Introdução ao problema de Sturm-Liouville

Valmir Peixôto de Souza Sobrinho

June 2021

Sumário

1 Revisão de Álgebra Linear

1.1 Espaços vetoriais

Seja A um conjunto não vazio. Dizemos que A é um espaço vetorial real se e somente se os axiomas abaixo forem satisfeitos.

Axioma 1. *Existe uma função* $S : A \times A \mapsto A, S(a,b) = c$ *na qual denominamos soma de elementos de* A*, ou simplesmente soma. Representamos a função* S *como* $a + b$*, a soma de* a *e* b*.*

Axioma 2. *Existe uma função* $M : \mathbb{R} \times A \mapsto A, M(a,x) = d$ *na qual denominamos multiplicação por números reais. Representamos a função* M *como* ax*, a multiplicação de* a *por* x*.*

Axioma 3. *A função soma é comutativa, ou seja,* $\forall a, b \in A : a + b = b + a$

Axioma 4. *A função soma é associativa, ou seja,* $\forall a, b, c \in A : (a + b) + c = a + (b + c)$

Axioma 5. *Existe um elemento* 0 *em* A *tal que* $a + 0 = a, \forall a \in A$

Axioma 6. *Para cada elemento* a *de* A*, existe um elemento* $-a$ *tal que* $a + (-a) = 0$

Axioma 7. $\forall a, b \in \mathbb{R}$ *e* $x \in A$*,* $a(bx) = (ab)x$

Axioma 8. $\forall a \in \mathbb{R}$ *e* $x, y \in A$*,* $a(x + y) = ax + ay$

Axioma 9. $\forall a, b \in \mathbb{R}$ *e* $x \in A$*,* $(a + b)x = ax + bx$

Axioma 10. $\forall x \in A, 1x = x$

Um espaço vetorial complexo possui os mesmos axiomas, com a diferença que os escalares são complexos.

Exemplo 1.1. *O conjunto de funções definidas em um intervalo, onde a soma de duas funções* f *e* g *é definida da forma usual* $f + g = f(x) + g(x), x \in I$ *e a multiplicação na forma usual* $af = af(x), a \in \mathbb{R}$*, é um espaço vetorial.*

Demonstração. Sejam f, g e h três funções definidas em um intervalo $I \subseteq \mathbb{R}$ e $a, b \in \mathbb{R}$.

- Uma vez que $f(x)$ e $g(x)$ retornam números reais, a soma entre eles comuta e ambos são definidos em I, então $f(x) + g(x) = g(x) + f(x)$
- A soma de números reais também é associativa, logo $(f(x)+g(x))+h(x) = f(x) + (g(x) + h(x))$
- Dada uma equação $f(x) + 0(x) = f(x)$, encontramos um $0(x) = 0, \forall x \in I$.
- Podemos obter uma função $(-f)(x)$ definida em I tal que $f(x)+(-f)(x) = 0$ fazendo $(-f)(x) = (-1)f(x) = -f(x)$, $\forall x \in I$

- Como $f(x) \in \mathbb{R}$ sempre que $x \in I$, então a multiplicação com números reais é associativa e $a(b \cdot f(x)) = (ab)f(x)$
- Sabendo que $f(x), g(x) \in \mathbb{R}$ em qualquer $x \in I$, então a propriedade distributiva dos números reais também se aplica em $a \cdot (f(x) + g(x)) = a \cdot f(x) + a \cdot g(x)$
- Similarmente a propriedade distributiva se aplica em $(a + b)f(x) = a \cdot f(x) + b \cdot f(x)$
- Como 1 é o elemento neutro multiplicativo dos reais, então $1 \cdot f(x) = f(x)$

■

Teorema 1.1 (Unicidade do elemento zero)**.** *Em qualquer espaço vetorial há apenas um único elemento zero.*

Demonstração. Seja A um espaço vetorial. Suponha que tenha pelo menos dois elementos zero, 0 e $0'$, em A. A soma $0 + 0'$ pode ser calculada de duas formas: aplicando o axioma do elemento zero de forma que $0'$ seja o elemento neutro aditivo da soma; ou aplicando esse axioma de maneira que 0 seja o elemento zero da soma.

Do primeiro jeito temos que $0 + 0' = 0$. Do outro, usando o axioma da comutatividade, obtemos $0 + 0' = 0' + 0 = 0'$

Portanto $0 = 0'$. Então todos os elementos zero de A são apenas um só. ■

Teorema 1.2 (Unicidade dos elementos simétricos)**.** *Em qualquer espaço vetorial, cada elemento x admite apenas um único simétrico $-x$.*

Demonstração. Sejam $-x$ e $(-x)'$ dois elementos simétricos de $x \in A$, um espaço vetorial. Usando a definição de simétrico e o axioma da soma comutativa temos:

$$x + (-x) = 0 \Longrightarrow x + (-x) + (-x)' = 0 + (-x)' \tag{1}$$

$$\Longrightarrow (x + (-x)') + (-x) = (-x)' \Longrightarrow -x = (-x)' \; . \tag{2}$$

Então todos os elementos simétricos de A são únicos. ■

Teorema 1.3. *Dado um espaço vetorial A, com $x, y \in A$ e $a, b \in \mathbb{R}$, então as propriedades abaixo são verdadeiras:*

1. $0x = 0$
2. $a0 = 0$
3. $(-a)x = -(ax) = a(-x)$
4. *Se $ax = 0$, logo $x = 0$ ou $a = 0$*
5. *Se $ax = ay$ e $a \neq 0$, logo $x = y$*

6. *Se* $ax = bx$ *e* $x \neq 0$*, logo* $a = b$

7. $-(x + y) = (-x) + (-y) = -x - y$

8. $\sum_{i=1}^{n} x = nx$

Demonstração. 1. $0x + 0x = (0 + 0)x = 0x$. Somando o simétrico de $0x$ (cuja existência em A é garantida pelo axioma (6) em ambos os lados da equação: $0x + 0x + (-0x) = 0x + (-0x) \implies 0x + 0 = 0 \implies 0x = 0$

2. $0 + 0 = 0 \implies a(0 + 0) = a0 \implies a0 + a0 = a0$. Somando o simétrico de $a0$ em ambos os lados da equação: $a0 + a0 + (-a0) = a0 + (-a0) \implies a0 + 0 = 0 \implies a0 = 0$

3. Primeiramente usamos a definição de simétrico de ax e depois somar $a(-x)$ de ambos os lados: $ax + (-(ax)) = 0 \implies ax + (-(ax)) + a(-x) = a(-x) \implies ax + a(-x) + (-(ax)) = a(-x)$

   E então: $a(x + (-x)) + (-(ax)) = a(-x) \implies a0 + (-(ax)) = a(-x) \implies 0 + (-(ax)) = a(-x) \implies -(ax) = a(-x)$

   Para a outra igualdade: $0 = 0x = (a + (-a))x = ax + (-a)x$, o que significa que $(-a)x$ é um elemento simétrico de ax. Mas $-(ax)$ também é um simétrico de ax, pois $ax + (-(ax)) = 0$. Como há apenas um único simétrico de ax, segue que $(-a)x = -(ax) = a(-x)$

4. Seja $ax = 0$. Se $a \neq 0$, então $\frac{1}{a} \cdot (ax) = \frac{1}{a} \cdot 0$, o que implica $(\frac{1}{a} \cdot a) \cdot x = 0 \implies 1x = 0 \implies x = 0$

   E se $a = 0$, então $ax = 0$ para qualquer elemento x de A.

5. Se $ax = ay$ e $a \neq 0$, então $\frac{1}{a} \cdot (ax) = \frac{1}{a} \cdot (ay)$ e $(\frac{1}{a} \cdot a)x = (\frac{1}{a} \cdot a)y \implies 1x = 1y \implies x = y$

6. Dado $ax = bx$ e $x \neq 0$, somamos o simétrico de bx em ambos os lados da equação, e então: $ax + (-bx) = bx + (-bx) \implies ax + (-bx) = 0 \implies (a - b)x = 0$. Usando a propriedade 4 e sabendo que $x \neq 0$, resta que $(a - b) = 0 \implies a = b$

7. Usando a propriedade 3 com $a = 1$, temos que $1(-y) = -(1y) = -y = (-1)y$. Então $(-x) + (-y) = -x + (-1)y = -x - (1y) = -x - y$. Além disso, $-(x + y) = (-1)(x + y) = (-1)x + (-1)y = -(1x) - (1y) = -x - y$

8. Primeiramente sabemos que $\sum_{i=1}^{1} x = 1x$. Suponha que $\sum_{i=1}^{n} x = nx$ é válido para algum algum número $n \in \mathbb{N}, n + 1 > 1$, como $n = 1$. Usando a definição de somatório com $n + 1 > 1$, temos que $\sum_{i=1}^{n+1} x = (\sum_{i=1}^{n} x) + x$, o que implica $\sum_{i=1}^{n+1} x = nx + x = (n + 1)x$. Então essa identidade está provada pelo princípio da indução $\forall n \in \mathbb{N}, n \geq 1$.

■

Dado um conjunto não-nulo $B \subseteq A$ e ambos A e B forem espaços vetoriais, então denominamos B como um subespaço vetorial de A.

Teorema 1.4. *Seja B um subconjunto de elementos do espaço vetorial A. B é um subespaço vetorial de A se e somente se a soma e multiplicação forem fechadas em B.*

Demonstração. Sejam $x, y \in B$. Sabemos que os axiomas (3), (4), (7), (8), (9) e (10) já são satisfeitos em B, pois $x, y \in B \implies x, y \in A$. Agora introduzimos a hipótese de que as operações de soma e multiplicação são fechadas em B, satisfazendo os axiomas (1) e (2). Sabemos também que $0x = 0$. Como a multiplicação é fechada, então segue que o elemento zero também pertence a B, satisfazendo o axioma 5. Analogamente, $(-1)x = -x$, o que implica $-x \in B$, satisfazendo o axioma 6.

E se B for um subespaço vetorial, então B é um espaço vetorial e portanto obedece todos os axiomas, incluindo o (1) e (2).

Portanto B é um subespaço vetorial de A se e somente se a soma e a multiplicação forem fechadas em B. ∎

Definição 1.1. *Seja x um elemento de um espaço vetorial A e S um subconjunto não-vazio de A. Se x pode ser escrito na forma*

$$x = \sum_{i=1}^{n} c_i x_i$$

onde $x_i, 1 \leq i \leq n$ são elementos do conjunto S e $c_i, 1 \leq i \leq n$ são escalares, então dizemos que x é uma combinação linear finita dos elementos de S.

Seja a um escalar e y uma outra combinação linear finita de S

$$y = \sum_{i=1}^{n} b_i x_i$$

Então

$$x + y = \sum_{i=1}^{n} c_i x_i + \sum_{i=1}^{n} b_i x_i = \sum_{i=1}^{n} c_i x_i + b_i x_i = \sum_{i=1}^{n} (c_i + b_i) x_i$$

$$ax = a \sum_{i=1}^{n} c_i x_i = \sum_{i=1}^{n} a(c_i x_i) = \sum_{i=1}^{n} (ac_i) x_i$$

Portanto as funções soma e multiplicação são fechadas no conjunto formado por todas as combinações lineares de S, denotado como $L(S)$, e portanto é um subespaço de A. Também chamamos $L(S)$ de subespaço gerado por S. Se S é o conjunto vazio, definimos $L(S) = 0$.

Definição 1.2. *Seja S um subconjunto de n elementos de A, $x_i \in S, c_i \in \mathbb{R}, 1 \leq i \leq n$. Se a relação abaixo*

$$\sum_{i=1}^{n} c_i x_i = 0$$

for verdadeira com pelo menos um coeficiente $c_i \neq 0$ então S é um conjunto linearmente dependente. Se esse for o caso, então a equação acima denota uma representação não-trivial do elemento zero.

Um subconjunto S com n elementos de A é linearmente independente se não for dependente, ou seja

$$\sum_{i=1}^{n} c_i x_i = 0 \implies \forall i \in \mathbb{N}, 1 \leq i \leq n, c_i = 0$$

Em outras palavras, a única solução possível para essa equação é a com todos os coeficientes nulos.

Corolário. *Se $0 \in S \subseteq A$, então S é linearmente dependente.*

Demonstração. Seja $S = \{x_i \in A, 1 \leq i \leq n, x_1 = 0\}$ Então a equação

$$\sum_{i=1}^{n} c_i x_i = c_1 x_1 + \sum_{i=2}^{n} c_i x_i = c_1 0 + \sum_{i=2}^{n} c_i x_i = 0$$

possui infinitas soluções não-triviais, como por exemplo $c_1 \in \mathbb{R}, \forall i : 2 \leq i \leq n : c_i = 0$. Logo S é linearmente dependente. ∎

Teorema 1.5. *Se S for um subconjunto linearmente independente com n elementos e $L(S)$ for o subespaço gerado por S, então qualquer subconjunto C com $n+1$ elementos de $L(S)$ é linearmente dependente.*

Demonstração. Se $n = 1$, temos $S = \{x_1\}$. Como S é linearmente independente, pela contrapositiva do corolário acima segue que $x_1 \neq 0$. Seja C um subconjunto com 2 elementos distintos de $L(S)$, $C = \{y_1, y_2\}$, ou seja $y_1 = c_1 x_1$ e $y_2 = c_2 x_1, c_1 \neq c_2$. Se $c_1 = 0$ ou $c_2 = 0$, então C é dependente. E se $c_1 \neq 0$ e $c_2 \neq 0$, podemos encontrar a_1 e a_2 não simultaneamente nulos tais que

$$a_1 y_1 + a_2 y_2 = 0 \implies a_1 c_1 x_1 + a_2 c_2 x_1 = 0 \implies (a_1 c_1 + a_2 c_2) x_1 = 0$$

Usando $a_1 = c_2$ e $a_2 = -c_1$, resulta que $a_1 \neq 0$ e $a_2 \neq 0$ e ao mesmo tempo a equação é satisfeita. Logo C é linearmente dependente.

Provado o teorema para $n = 1$, usamos o princípio da indução para provar que também é válido para $n \geq 1$.

Suponha que um subconjunto C que tem $n+1$ elementos distintos de $L(S)$, com S contendo n elementos, implica que C é linearmente dependente. Devemos provar que se S tem $n+1$ elementos e C tem $n+2$ elementos de $L(S)$, então C é dependente.

Sejam os elementos de $C = \{y_i, 1 \leq i \leq n+2\}$. Como $C \subseteq L(S)$, segue que cada elemento de C pode ser escrito como uma combinação linear dos elementos de S:

$$y_i = \sum_{j=1}^{n+1} a_{ij} x_j$$

Se existir algum $a_{kl} \neq 0$ para um certo l, então

$$y_k = a_{kl}x_l + \sum_{j=1, j\neq l}^{n+1} a_{kj}x_j$$

$$\frac{a_{il}}{a_{kl}}y_k = a_{il}x_l + \sum_{j=1, j\neq l}^{n+1} \frac{a_{il}a_{kj}}{a_{kl}}x_j$$

Subtraindo essa equação pela combinação linear de y_i, temos

$$\frac{a_{il}}{a_{kl}}y_k - y_i = \sum_{j=1, j\neq l}^{n+1} (\frac{a_{il}a_{kj}}{a_{kl}} - a_{ij})x_j \quad (3)$$

Isso significa que $\frac{a_{il}}{a_{kl}}y_k - y_i$ pode ser escrito como uma combinação linear de $S' = \{x_j, 1 \leq j \leq n+1, j \neq k\}$ que é um conjunto com n elementos. Como S é independente, então $S' \subseteq S$ também é independente. A equação (3) com $i = k$ não nos fornece nada útil. Então há $n+1$ termos $\frac{a_{il}}{a_{kl}}y_k - y_i$ pertencentes a $L(S')$. O conjunto destes então, por hipótese, é linearmente dependente. Logo há escalares b_i não todos nulos tais que

$$\sum_{i=1, i\neq k}^{n+2} b_i(\frac{a_{il}}{a_{kl}}y_k - y_i) = 0$$

implica

$$\sum_{i=1, i\neq k}^{n+2} b_i\frac{a_{il}}{a_{kl}}y_k - \sum_{i=1, i\neq k}^{n+2} b_iy_i = 0$$

$$(\sum_{i=1, i\neq k}^{n+2} b_i\frac{a_{il}}{a_{kl}})y_k - \sum_{i=1, i\neq k}^{n+2} b_iy_i = 0$$

Essa é uma representação não-trivial do elemento zero pelos elementos de C, logo C é linearmente dependente.

E se não existir nenhum $a_{kl} \neq 0$ para um certo l, ou seja, $a_{kl} = 0, \forall k, 1 \leq k \leq n+1$, então

$$y_i = \sum_{j=1, j\neq l}^{n+1} a_{ij}x_j$$

Como o conjunto $S' = \{x_j, 1 \leq j \leq n+1, j \neq k\}$ tem n elementos, por hipótese, qualquer conjunto com $n+1$ elementos ou mais é linearmente dependente, o que é o caso de C com $n+2$ elementos, o que completa a prova.

■

1.2 Bases e dimensão de espaços vetoriais

Definição 1.3. *Dado um subconjunto B de finitos elementos de um espaço vetorial de A. Se B é linearmente independente e $L(B) = A$, então denominamos B como uma base finita de A. Se A possui uma base finita ou for constituído apenas pelo elemento nulo, então dizemos que A tem dimensão finita. E se A não tiver uma base finita, nem $A = \{0\}$, então A tem dimensão infinita.*

Teorema 1.6. *Se o espaço vetorial A tem dimensão finita, então todas as bases finitas de A terão a mesma quantidade de elementos.*

Demonstração. Seja um subconjunto B uma base finita de A e que possui n elementos. Seja também um subconjunto C de A. Se C possui m elementos, $m > n$, então o teorema (1.5) garante que C é linearmente dependente, já que $L(B) = A$ e B tem n elementos. Logo C não pode ser uma base finita se tiver mais do que n elementos.

Então suponha que C tem m elementos, $m < n$, e que $L(C) = A$ é verdade. Isso implica que B não pode ser uma base finita por ser linearmente dependente, também pelo teorema (1.5), o que contradiz a hipótese. Ou seja, $L(C) \neq A$, e portanto C não pode ser uma base finita. Resta que se C é uma base finita, então C tem exatamente n elementos. ■

Definição 1.4. *Se um espaço vetorial A tem uma base finita com n elementos, então dizemos que A tem dimensão n, ou $dimA = n$. E se $A = \{0\}$, então $dimA = 0$*

Teorema 1.7. *Seja A um espaço vetorial com dimensão finita e $dimA = n$. Logo:*

- *Qualquer subconjunto linearmente independente de A é também um subconjunto de alguma base de A.*
- *Qualquer subconjunto linearmente independente de A com n elementos é uma base de A.*

Demonstração. Primeira afirmação: seja B um conjunto linearmente independente de elementos de A. Se $L(B) = A$, então B é uma base. Caso contrário, existe algum elemento de A que não pertence ao conjunto de todas as combinações lineares dos elementos de B. Seja $x_1 \in A, x_1 \notin L(B)$. Seja também $C = B \cup \{x_1\}$. Então C é linearmente independente. Se $L(C) = A$, então C é uma base de A na qual $B \subseteq C$. Caso contrário, existirá algum outro elemento $x_2 \in A, x_2 \notin L(C)$. Como B nesse caso tem menos elementos que uma base de A (uma vez que B é linearmente independente), repetindo o processo de inserir um elemento independente resultará inexoravelmente em um conjunto linearmente independente D com n elementos no qual B é subconjunto. Se $L(D) = A$, então D é uma base de A. E se $L(D) \neq A$, isso significa que existe algum outro elemento independente em A, x_{n+1}, o que significaria que $D \cup \{x_{n+1}\}$ é linearmente independente, o que é um absurdo, pois qualquer conjunto com $n + 1$

elementos de A é dependente, conforme o Teorema (1.5). Isso significa que D obrigatoriamente gera A, e portanto B é subconjunto de alguma base D de A.

Segunda afirmação: se um conjunto independente D tem n elementos, então D é um subconjunto de alguma base B. Como $dimA = n$, segue que B tem n elementos. Como ambos os conjuntos possuem a mesma quantidade de elementos e $D \subseteq B$, segue que $D = B$ e portanto D é uma base de A. ■

Corolário. *Dado uma base finita $B = \{x_1, x_2, x_3, ..., x_n\}$ de A, $dimA = n$ e $x \in A$, então só existe uma única solução dos escalares c_i para a equação*

$$x = \sum_{i=1}^{n} c_i x_i$$

Demonstração. Com efeito, se x também puder ser representado por outros escalares d_i na mesma base B, então

$$x = \sum_{i=1}^{n} d_i x_i = \sum_{i=1}^{n} c_i x_i$$

e logo

$$\sum_{i=1}^{n} d_i x_i - c_i x_i = 0$$

$$\sum_{i=1}^{n} (d_i - c_i) x_i = 0 \tag{4}$$

Como o conjunto B é linearmente independente, segue que

$$d_i - c_i = 0$$

$$c_i = d_i, \forall i, 1 \leq i \leq n \tag{5}$$

Então a solução dos escalares é única para uma certa base B.

■

1.2.1 Componentes

Dado uma base finita $B = \{e_i, 1 \leq i \leq n\}$ de um espaço vetorial A de dimensão n e $x \in A$ então

$$x = \sum_{i=1}^{n} c_i e_i$$

para alguns escalares c_i. Podemos associar esses escalares a uma n-upla ordenada do $\mathbb{R}^n$, $(c_1, c_2, c_3, ..., c_n)$, e a base B à $(e_1, e_2, e_3, ..., e_n)$ do $\mathbb{R}^n$. Como o corolário do teorema (1.7) garante que só existe uma única solução dos escalares

c_i para representar x na base B, então segue que o elemento x pode ser biunivocamente relacionado com a n-upla ordenada dos seus coeficientes sob uma determinada base. Então dizemos que a n-upla ordenada $(c_1, c_2, c_3, ..., c_n)$ tem as coordenadas do elemento x relativamente a base ordenada $(e_1, e_2, e_3, ..., e_n)$.

Note que se representarmos a base B em termos da base B, então obtemos

$$e_i = \sum_{j=1}^{n} \delta_{ij} e_i$$

onde

$$\delta_{ij} = \begin{cases} 0, & \text{se} \quad i \neq j \\ 1, & \text{se} \quad i = j \end{cases}$$

o que implica que e_i está relacionado à n-upla $(\delta_{i1}, \delta_{i2}, \delta_{i3}, ..., \delta_{ii}, ..., \delta_{in})$, no que resulta em uma n-upla com um 1 no índice i e $n-1$ zeros nos índices restantes.

1.3 Produto interno e normas

Definição 1.5. *Seja A um espaço vetorial, $x, y, w \in A$ e $z, m \in \mathbb{C}$. Seja uma função $P : A \times A \longmapsto \mathbb{C}, P(x, y) = z$. Dizemos que P é um produto interno se e somente se satisfazer os seguintes axiomas:*

Axioma 11 (Comutatividade)**.** $P(x, y) = \overline{P(y, x)}$

Axioma 12 (Distribuição)**.** $P(x, y + w) = P(x, y) + P(x, w)$

Axioma 13 (Multiplicação)**.** $mP(x, y) = P(x, my)$

Axioma 14 (Positividade)**.** $P(x, x) > 0$ *se* $x \neq 0$

Note que $P(x, x) = \overline{P(x, x)}$, ou seja, $P(x, x) \in \mathbb{R}$.

Corolário. $mP(x, y) = P(\overline{m}x, y)$

Demonstração. $mP(x, y) = m\overline{P(y, x)} = \overline{P(y, \overline{m}x)} = P(\overline{m}x, y)$ ∎

Corolário. $P(0, 0) = 0$

Demonstração. $P(0, 0) = P(0, 0 + 0) = P(0, 0) + P(0, 0)$. Subtraindo $P(0, 0)$ de ambos os lados obtemos $0 = P(0, 0)$

∎

Corolário. $P(0, y) = 0$

Demonstração. $P(0, y) = P(0 \cdot 0, y) = 0P(0, y) = 0$ ∎

Denotamos o produto interno $P(x, y) = z$ como $\langle x, y \rangle = z$.

Se um espaço vetorial complexo A é provido de produto interno, então chamamos A de espaço unitário. Se A for um espaço vetorial real munido de produto interno, chamamos A de espaço euclidiano.

Exemplo 1.2. *Seja A o espaço vetorial das funções complexas contínuas de uma variável definidas em um intervalo $I = [a, b]$, $w(x)$ uma função positiva contínua em I e $f, g, h \in A$. A função*

$$P(f, g) = \langle f, g \rangle = \int_a^b w(x) f(x) \overline{g(x)} dx \tag{6}$$

é um produto interno em A.

Demonstração.

$$\begin{aligned} \langle f, g \rangle &= \int_a^b w(x) f(x) \overline{g(x)} dx \\ &= \int_a^b w(x) \overline{g(x) \overline{f(x)}} dx \\ &= \overline{\int_a^b w(x) g(x) \overline{f(x)} dx} \\ &= \overline{\langle g, f \rangle} \end{aligned}$$

$$\begin{aligned} \langle f, g + h \rangle &= \int_a^b w(x) f(x) \cdot (\overline{g(x) + h(x)}) dx \\ &= \int_a^b w(x) f(x) \overline{g(x)} + w(x) f(x) \overline{h(x)} dx \\ &= \int_a^b w(x) f(x) \overline{g(x)} dx + \int_a^b w(x) f(x) \overline{h(x)} dx \\ &= \langle f, g \rangle + \langle f, h \rangle \end{aligned}$$

Seja $z \in \mathbb{C}$. Então

$$z \langle f, g \rangle = z \int_a^b w(x) f(x) \overline{g(x)} dx = \int_a^b w(x) f(x) (m \cdot \overline{g(x)}) dx = \langle f, mg \rangle$$

$$\langle f, f \rangle = \int_a^b w(x) f(x) \overline{f(x)} dx = \int_a^b w(x) |f(x)|^2 dx$$

Como $w(x) > 0$ e $f^2(x)$ é contínuo e não-negativo no intervalo I, então $w(x) f^2(x) \geq 0$ e podemos usar a propriedade das integrais positivas:

$$\langle f, f \rangle = \int_a^b w(x) |f(x)|^2 dx \geq 0$$

Como o integrando é não-negativo, segue que a integral só é zero se $w(x) = 0$ ou $f(x) = 0$. Sabendo que $w(x) > 0$ por definição, segue que $\langle f, f \rangle$ só é zero se $f(x)$ for o elemento nulo. E se $f(x) \neq 0$, temos

$$f(x) \neq 0 \implies \langle f, f \rangle = \int_a^b w(x) |f(x)|^2 dx > 0$$

■

Teorema 1.8. *[Desigualdade de Cauchy-Schwarz] Dado um espaço unitário A, segue que*

$$\forall x, y \in A : |\langle x, y\rangle|^2 \leq \langle x, x\rangle \cdot \langle y, y\rangle$$

Demonstração. Sejam a e b dois escalares. Então

$$\begin{aligned} \langle ax + by, ax + by\rangle &= \langle ax + by, ax\rangle + \langle ax + by, by\rangle \\ &= \langle ax, ax\rangle + \langle by, ax\rangle + \langle ax, by\rangle + \langle by, by\rangle \\ &= a\overline{a}\,\langle x, x\rangle + b\overline{a}\,\langle y, x\rangle + a\overline{b}\,\langle x, y\rangle + b\overline{b}\,\langle y, y\rangle \geq 0 \end{aligned} \tag{7}$$

Fazendo $a = \langle y, y\rangle$ e $b = -\langle x, y\rangle$, então $\overline{a} = \overline{\langle y, y\rangle} = \langle y, y\rangle$, $\overline{b} = -\overline{\langle x, y\rangle} = -\langle y, x\rangle$ e dividindo a inequação acima por $\langle y, y\rangle$ temos

$$\begin{aligned} \langle y, y\rangle\, \langle x, x\rangle - \langle y, x\rangle\, \langle x, y\rangle - \langle x, y\rangle\, \langle y, x\rangle + (-\langle x, y\rangle) \cdot (-\langle y, x\rangle) \geq 0 \\ \langle y, y\rangle\, \langle x, x\rangle + [-\langle y, x\rangle\, \langle x, y\rangle + \langle y, x\rangle\, \langle x, y\rangle] - |\langle x, y\rangle|^2 \geq 0 \\ \langle y, y\rangle\, \langle x, x\rangle \geq |\langle x, y\rangle|^2 \end{aligned}$$

■

Corolário. *A igualdade* $|\langle x, y\rangle|^2 = \langle y, y\rangle\, \langle x, x\rangle$ *só é válida se e somente se os elementos* x *e* y *são linearmente dependentes.*

Demonstração. Se x e y são linearmente dependentes, então $x = cy$ para algum escalar c e

$$\begin{aligned} \langle y, y\rangle\, \langle x, x\rangle &= \langle y, y\rangle\, \langle cy, cy\rangle \\ &= c\overline{c}\,\langle y, y\rangle\, \overline{\langle y, y\rangle} \\ &= [c \cdot \langle y, y\rangle] \cdot \overline{[c \cdot \langle y, y\rangle]} \\ &= |c\,\langle y, y\rangle|^2 \\ &= |\langle y, cy\rangle|^2 \\ &= |\langle y, x\rangle|^2 \\ &= |\langle x, y\rangle|^2 \quad \text{(Módulo não altera com o conjugado)} \end{aligned}$$

Por outro lado, se $|\langle x, y\rangle|^2 = \langle y, y\rangle\, \langle x, x\rangle$, então há duas possibilidades: x ou y é o elemento nulo, o que torna x e y dependentes; ou então nenhum dos dois é o elemento nulo. Neste caso, substituindo a desigualdade na demonstração do teorema de Cauchy-Schwarz (8):

$$\forall x, y \in A : |\langle x, y\rangle|^2 \leq \langle x, x\rangle \cdot \langle y, y\rangle \tag{8}$$

pela igualdade em (7), concluímos que

$$\langle ax + by, ax + by \rangle = 0$$

Pelo axioma (14), isso só é possível se

$$ax + by = 0$$

Mas também sabemos que $a = \langle y, y \rangle$ e $b = -\langle x, y \rangle$. Então

$$ax + by = \langle y, y \rangle \cdot x - \langle x, y \rangle \cdot y = 0$$

Por hipótese, $y \neq 0$, o que implica em $\langle y, y \rangle \neq 0$. Isso significa que encontramos uma solução não-trivial para a combinação linear de x e y igual a zero. Logo x e y são dependentes.

■

Note que como $\langle y, x \rangle = \overline{\langle x, y \rangle}$, segue que a soma

$$\langle x, y \rangle + \overline{\langle x, y \rangle} = \langle x, y \rangle + \langle y, x \rangle$$

pertence aos reais e também

$$\langle x, y \rangle + \langle y, x \rangle = 2 \operatorname{Re} \langle x, y \rangle \tag{9}$$

Isso também significa

$$\begin{aligned} \|x - y\|^2 &= \langle x - y, x - y \rangle \\ &= \|x\|^2 - \langle x, y \rangle - \langle y, x \rangle + \|y\|^2 \\ &= \|x\|^2 - 2 \operatorname{Re} \langle x, y \rangle + \|y\|^2 \end{aligned} \tag{10}$$

1.3.1 Norma

Definição 1.6. *Dado um espaço unitário A, definimos a norma $\|x\|$ de um elemento $x \in A$ pela equação* (11)*:*

$$\|x\| = \sqrt{\langle x, x \rangle} \tag{11}$$

A desigualdade de Cauchy-Schwarz pode ser reescrita como

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\| \tag{12}$$

Teorema 1.9. *Seja A um espaço unitário, m um escalar e $x, y \in A$. Então:*

1. $\|x\| = 0$ *se e somente se* $x = 0$
2. $\|x\| > 0$ *se* $x \neq 0$ *(Positividade)*
3. $\|cx\| = |c| \cdot \|x\|$
4. $\|x + y\| \leq \|x\| + \|y\|$ *(Desigualdade triangular)*

Demonstração. 1. Se $x = 0$, logo $\|x\| = \|0\| = \sqrt{\langle 0, 0 \rangle} = 0$

E se $\|x\| = 0$, então $\|x\| = \sqrt{\langle x, x \rangle} = 0$, o que implica $\langle x, x \rangle = 0$. Pelo axioma (14), temos que $\langle x, x \rangle = 0$ só é possível se $x = 0$.

2. $\|x\| = \sqrt{\langle x, x \rangle}$. Como $\langle x, x \rangle > 0$ se $x \neq 0$ (axioma (14)), então $\sqrt{\langle x, x \rangle} > 0$ e $\|x\| > 0$ se $x \neq 0$.

3. $\|cx\| = \sqrt{\langle cx, cx \rangle} = \sqrt{c\bar{c} \langle x, x \rangle} = \sqrt{|c|^2 \langle x, x \rangle} = |c|\sqrt{\langle x, x \rangle} = |c| \cdot \|x\|$

4. $||x + y||^2 = \langle x + y, x + y \rangle = \langle x + y, x \rangle + \langle x + y, y \rangle = \langle x, x \rangle + \langle y, x \rangle + \langle x, y \rangle + \langle y, y \rangle = \|x\|^2 + \|y\|^2 + \langle y, x \rangle + \langle x, y \rangle$. Pela desigualdade de Cauchy-Schwarz, $| \langle y, x \rangle | \leq \|y\|\|x\|$ e $| \langle x, y \rangle | \leq \|x\|\|y\|$. Como $\langle y, x \rangle = \overline{\langle x, y \rangle}$, segue que $\langle x, y \rangle + \langle y, x \rangle$ é um número pertencente aos reais, que é um corpo ordenado. Portanto, juntando as duas inequações de Cauchy-Schwarz, temos $\langle x, y \rangle + \langle y, x \rangle \leq \|x\|\|y\| + \|x\|\|y\| = 2\|x\|\|y\|$. Substituindo a soma na equação do quadrado da norma obtemos $\|x + y\|^2 = \|x\|^2 + \|y\|^2 + \langle y, x \rangle + \langle x, y \rangle \leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| = (\|x\| + \|y\|)^2$. Aplicando a raíz quadrada ao resultado resulta em $\|x + y\| \leq \|x\| + \|y\|$

■

Definição 1.7. *Dado um espaço euclidiano A, o ângulo θ entre dois elementos não-nulos $x, y \in A : x \neq 0, y \neq 0$ é definido da seguinte forma:*

$$\cos \theta = \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|} \quad \theta \in [0, \pi]$$

Uma vez que $| \langle x, y \rangle | \leq \|x\| \cdot \|y\|$, segue então que

$$| \langle x, y \rangle | \leq \|x\| \cdot \|y\| \implies \left| \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|} \right| \leq 1 \implies -1 \leq \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|} \leq 1$$

Sabendo que a função cosseno é bijetora no intervalo $[0, \pi]$ na imagem $[-1, 1]$, então conclui-se que θ é único para cada par de elementos não-nulos de A.

Teorema 1.10. *Dado um espaço euclidiano A, a norma de um elemento $x \in A$ pode ser dado pela expressão*

$$\|x\| = \sup_{\|y\|=1} | \langle y, x \rangle |$$

Demonstração. Se $x = 0$, rapidamente temos $0 = \sup_{\|y\|=1} | \langle y, 0 \rangle | = 0$.

Agora suponha $x \neq 0$. Seja $y = \dfrac{x}{\|x\|}$. Então

$$\begin{aligned} \|x\|^2 &= \langle x, x \rangle \\ &= \|x\| \cdot \left\langle \frac{x}{\|x\|}, x \right\rangle \\ &= \|x\| \cdot \langle y, x \rangle \\ &\leq \|x\| \cdot \sup_{\|y\|=1} |y, x| \end{aligned}$$

ou seja,

$$\|x\| \leq \sup_{\|y\|=1} |\langle y, x\rangle| \tag{13}$$

Por outro lado, a desigualdade de Cauchy-Schwarz nos diz que

$$|\langle y, x\rangle| \leq \|y\| \cdot \|x\| \leq \|x\|$$

e isso inclui

$$\sup_{\|y\|=1} |\langle y, x\rangle| \leq \|x\| \tag{14}$$

Juntando as duas inequações (13) e (14), temos

$$\|x\| = \sup_{\|y\|=1} |\langle y, x\rangle|$$

■

1.3.2 Ortogonalidade

Definição 1.8. *Em um espaço unitário A, dizemos que $x \in A$ e $y \in A$ são ortogonais se e somente se $\langle x, y\rangle = 0$. Um subconjunto $B \subseteq A$ se diz ortogonal se $\forall x, y \in B$, $\langle x, y\rangle = 0$. Um conjunto ortogonal C pode ser chamado de ortonormal se $\forall x \in C, \|x\| = 1$*

O axioma (14) garante que se $x \in A : \langle x, x\rangle = 0 \implies x = 0$, o que significa que o elemento nulo é o único elemento ortogonal a ele próprio.

Além disso, o corolário do axioma (13) garante que $\langle 0, x\rangle = 0$, ou seja, todos os elementos do espaço unitário A são ortogonais ao elemento nulo.

Teorema 1.11. *Qualquer conjunto ortogonal B de elementos não-nulos de A é linearmente independente. E se A tiver dimensão finita, $dimA = n$ e B tem n elementos, então B é uma base de A.*

Demonstração. Sejam $x_i \in B$, $1 \leq i \leq n$ os únicos elementos de B e possuem a propriedade $\langle x_i, x_j\rangle = 0$ se $i \neq j$. Dado a equação

$$\sum_{i=1}^{n} c_i x_i = 0$$

com os coeficientes c_i desconhecidos, podemos aplicar o produto interno nesse somatório com qualquer elemento x_m de B:

$$\langle \sum_{i=1}^{n} c_i x_i, x_m\rangle = 0$$

Usando os axiomas (12) e (13), obtemos

$$\sum_{i=1}^{n} \overline{c_i} \langle x_i, x_m\rangle = 0$$

Como os elementos são ortogonais entre si, então só restará 1 termo não-nulo do somatório, o termo em que $i = m$, e então

$$\overline{c_m} \cdot \langle x_m, x_m \rangle = 0$$

Por definição, nenhum dos elementos de B é o elemento nulo, o que implica em $\langle x_m, x_m \rangle \neq 0$, o que resta $\overline{c_m} = 0 \implies c_m = 0$. Varrendo m de 1 a n obteremos $\forall i : 1 \leq i \leq n : c_i = 0$, o que implica que B é linearmente independente.

Se $dimA = n$, n o número de elementos de B, então, pelo teorema (1.7) , B forma uma base de A. ■

Exemplo 1.3. *Seja A o espaço unitário das funções de uma variável definidas no intervalo $[0, 2\pi]$ e o produto interno definido no exemplo 1.2 com $a = 0$ e $b = 2\pi$. Seja também o conjunto $B \subseteq A$ formado pelos elementos*

$$u_0(x) = 1 \qquad u_{2n-1}(x) = cos(nx) \quad u_{2n}(x) = sin(nx) \quad , \textit{para } n = 1, 2, 3, ...$$

Se $m \neq n$, então

$$\int_0^{2\pi} u_n(x) u_m(x) dx = 0$$

Demonstração. Sejam m e n números inteiros distintos maiores do que zero. Como m e n são distintos, então $2m$ e $2n$ são números pares distintos e

$$u_{2n}(x) = sin(nx) \qquad u_{2m}(x) = sin(mx)$$

Seja I_1 o produto interno $\langle u_{2n}(x), u_{2m}(x) \rangle$. Realizando integração por partes duas vezes, obtemos:

$$I_1 = \int_0^{2\pi} sin(nx) sin(mx) dx \tag{15}$$

$$= \underbrace{-sin(nx)\frac{cos(mx)}{m}\Bigg|_0^{2\pi}}_{=0} \overbrace{- \int_0^{2\pi} -\frac{n}{m} cos(nx) cos(mx) dx}^{=I_2} \tag{16}$$

$$= \int_0^{2\pi} n \cdot cos(nx) \frac{cos(mx)}{m} dx \tag{17}$$

$$= \underbrace{-n \cdot cos(nx) \frac{sin(mx)}{m^2}\Bigg|_0^{2\pi}}_{=0} - \int_0^{2\pi} -\frac{n^2}{m^2} sin(nx) sin(mx) dx \tag{18}$$

$$= \frac{n^2}{m^2} \cdot I_1 \tag{19}$$

Portanto temos $I_1 = \frac{n^2}{m^2} \cdot I_1$. Como m e n são diferentes por hipótese, então segue que $I_1 = 0$.

Como n e m são números distintos, então $2n-1$ e $2m-1$ são dois números ímpares distintos e

$$u_{2n-1}(x) = cos(nx) \qquad u_{2m-1}(x) = cos(mx)$$

Seja I_2 o produto interno $\langle u_{2n-1}(x), u_{2m-1}(x) \rangle$. Observando a equação (16) de I_1 e considerando que $I_1 = 0$ se $n \neq m$ então:

$$I_2 = \int_0^{2\pi} cos(nx)cos(mx)dx = I_1 \cdot \frac{m}{n} = 0$$

Continuando, $2n$ e $2m-1$ são números distintos, uma vez que um é par e outro é ímpar, então:

$$u_{2n}(x) = sin(nx) \qquad u_{2m-1}(x) = cos(mx)$$

Seja I_3 o produto interno $\langle u_{2n}(x), u_{2m-1}(x) \rangle$. Realizando integração por partes duas vezes obtemos

$$\begin{aligned}
I_3 &= \int_0^{2\pi} sin(nx)cos(mx)dx \\
&= \int_0^{2\pi} sin(nx)cos(mx)dx \\
&= \underbrace{sin(nx)\frac{sin(mx)}{m}\Bigg|_0^{2\pi}}_{=0} - \int_0^{2\pi} n \cdot cos(nx)\frac{sin(mx)}{m}dx \\
&= (-1) \cdot \left[\underbrace{-n \cdot cos(nx)\frac{cos(mx)}{m^2}\Bigg|_0^{2\pi}}_{=0} - \int_0^{2\pi} \frac{n^2}{m^2} sin(nx)cos(mx)dx \right] \\
&= \frac{n^2}{m^2} \int_0^{2\pi} sin(nx)cos(mx) \\
&= \frac{n^2}{m^2} \cdot I_3
\end{aligned}$$

$$\frac{n^2}{m^2} \neq 1 \implies I_3 = 0$$

Por fim, se $n = 0$ e $m \neq 0$, temos os índices possíveis são 0 e $2m$ ou 0 e $2m-1$.

No primeiro caso temos

$$u_0(x) = 1 \qquad u_{2m}(x) = sin(mx)$$

Seja I_4 o produto interno $\langle u_0(x), u_{2m}(x) \rangle$. Então

$$I_4 = \int_0^{2\pi} 1 \cdot sin(mx)dx = -\frac{cos(mx)}{m}\Bigg|_0^{2\pi} = \frac{cos(mx)}{m}\Bigg|_{2\pi}^{0} = \frac{1}{m} \cdot (1-1) = 0$$

No segundo caso:

$$u_0(x) = 1 \qquad u_{2m-1}(x) = cos(mx)$$

Seja I_5 o produto interno $\langle u_0(x), u_{2m-1}(x) \rangle$. Logo

$$I_5 = \int_0^{2\pi} 1 \cdot cos(mx)dx = \frac{sin(mx)}{m}\Bigg|_0^{2\pi} = 0$$

Portanto, dado quaisquer dois inteiros não-negativos distintos n' e m', independente de serem par ou ímpar, o produto interno entre $u_{n'}$ e $u_{m'}$ é igual a 0.

■

Isso significa que todos os elementos do conjunto B são ortogonais dois a dois e B é, então, um conjunto ortogonal. Visto que a função nula não pertence ao conjunto ortogonal B então, pelo teorema 1.10, B é linearmente independente.

Teorema 1.12. *Seja A um espaço unitário de dimensão finita n que possui uma base ortogonal $B = \{e_1, e_2, e_3, ..., e_n\}$. Seja também um elemento $x \in A$ que pode ser escrito como uma combinação linear de B:*

$$x = \sum_{i=1}^{n} c_i e_i$$

Então todos os escalares c_i, $1 \leq i \leq n$ podem ser encontrados usando o produto interno:

$$c_i = \frac{\langle e_i, x \rangle}{\langle e_i, e_i \rangle}$$

Demonstração. Seja o termo $\langle e_i, x \rangle$ o produto interno entre o elemento de índice i da base B e o elemento x. Então:

$$\langle e_i, x \rangle = \langle e_i, \sum_{k=1}^{n} c_k e_k \rangle = \sum_{k=1}^{n} \langle e_i, c_k e_k \rangle = \underbrace{\sum_{k=1}^{n} c_k \langle e_k, e_i \rangle = c_i \langle e_i, e_i \rangle}_{\langle e_k, e_i \rangle = 0 \iff k \neq i}$$

onde a somatória na última expressão foi simplificada sabendo que B é um conjunto ortogonal. Obtemos então:

$$\langle e_i, x \rangle = c_i \langle e_i, e_i \rangle$$

$$c_i = \frac{\langle e_i, x \rangle}{\langle e_i, e_i \rangle}$$

■

Se $x, y \in A$ são elementos cujos coeficientes já são conhecidos e há uma base ortogonal $B = \{e_1, e_2, e_3, ..., e_n\}$, então

$$\begin{aligned} \langle x, y \rangle &= \langle \sum_{i=1}^{n} c_i e_i, \sum_{i=1}^{n} d_i e_i \rangle \\ &= \sum_{k=1}^{n} \langle c_k e_k, \sum_{i=1}^{n} d_i e_i \rangle \\ &= \sum_{k=1}^{n} \sum_{m=1}^{n} \langle c_k e_k, d_m e_m \rangle \\ &= \sum_{k=1}^{n} \sum_{m=1}^{n} d_k \overline{c_m} \langle e_k, e_m \rangle \\ &= \sum_{i=1}^{n} d_i \overline{c_i} \langle e_i, e_i \rangle \end{aligned} \tag{20}$$

Se a base ortogonal B também for uma base normalizada, ou seja, $\forall e_i \in B : ||e_i|| = 1$, o que implica $\langle e_i, e_i \rangle = ||e_i||^2 = 1$, então

$$\langle x, y \rangle = \sum_{i=1}^{n} d_i \overline{c_i} \tag{21}$$

Isso significa que se A for um espaço real (onde $\overline{c_i} = c_i$), então podemos calcular qualquer produto interno entre 2 elementos de um espaço unitário qualquer fazendo o produto escalar entre as n-uplas dos coeficientes desses elementos relativos a uma base ortonormal qualquer.

Uma vez que a n-upla ordenada dos escalares que representam o elemento x pode ser relacionada a um único elemento do $\mathbb{R}^n$ (o mesmo para y), segue que podemos considerar apenas a n-upla dos coeficientes, relativos a uma base ortonormal, para calcular o produto interno, independentemente de como este foi definido.

Teorema 1.13. *Seja A um espaço unitário de dimensão finita n que possui uma base ortonormal $B = \{e_1, e_2, e_3, ..., e_n\}$. Então $\forall x, y \in A$:*

$$\langle x, y \rangle = \sum_{i=1}^{n} \langle e_i, y \rangle \, \overline{\langle e_i, x \rangle}$$

E no caso particular $x = y$:

$$\|x\|^2 = \sum_{i=1}^{n} | \langle e_i, x \rangle |^2$$

Demonstração. Conforme a equação (21), se B é ortonormal, c_i são as componentes de x e d_i as de y, ambas em relação a B, segue que

$$\langle x, y \rangle = \sum_{i=1}^{n} d_i \overline{c_i} = \sum_{i=1}^{n} \frac{\langle e_i, y \rangle}{\langle e_i, e_i \rangle} \cdot \frac{\overline{\langle e_i, x \rangle}}{\overline{\langle e_i, e_i \rangle}} = \sum_{i=1}^{n} \langle e_i, y \rangle \, \overline{\langle e_i, x \rangle}$$

■

Teorema 1.14 (Desigualdade de Bessel)**.** *Seja A um espaço unitário de dimensão infinita e B um conjunto infinito de elementos ortonormais $\{e_1, e_2, e_3, \ldots\}$. Então $\forall x \in A$:*

$$\sum_{i=1}^{\infty} | \langle e_i, x \rangle |^2 \leq \|x\|^2$$

Demonstração. Do axioma (14) e da equação (10), é válido $\forall n \geq 1$:

$$\begin{aligned} 0 \leq \left\| x - \sum_{i=1}^{n} \langle e_i, x \rangle \, e_i \right\|^2 &= \left\langle x - \sum_{i=1}^{n} \langle e_i, x \rangle \, e_i, \; x - \sum_{i=1}^{n} \langle e_i, x \rangle \, e_i \right\rangle \\ &= \|x\|^2 - 2 \operatorname{Re} \left\langle x, \sum_{i=1}^{n} \langle e_i, x \rangle \, e_i \right\rangle + \left\| \sum_{i=1}^{n} \langle e_i, x \rangle \, e_i \right\|^2 \end{aligned}$$

A última norma pode ser calculada usando a equação (21):

$$\begin{aligned} \left\| \sum_{i=1}^{n} \langle e_i, x \rangle \, e_i \right\|^2 &= \left\langle \sum_{i=1}^{n} \langle e_i, x \rangle \, e_i, \; \sum_{i=1}^{n} \langle e_i, x \rangle \, e_i \right\rangle \\ &= \sum_{i=1}^{n} \langle e_i, x \rangle \, \overline{\langle e_i, x \rangle} \\ &= \sum_{i=1}^{n} | \langle e_i, x \rangle |^2 \end{aligned} \tag{22}$$

e portanto

$$
\begin{aligned}
0 &\leq \|x\|^2 - 2\,\mathrm{Re}\left\langle x, \sum_{i=1}^{n} \langle e_i, x\rangle\, e_i \right\rangle + \sum_{i=1}^{n} |\, \langle e_i, x\rangle|^2 \\
&\leq \|x\|^2 - 2\sum_{i=1}^{n} \mathrm{Re}\ \langle x,\ \langle e_i, x\rangle\, e_i\rangle + \sum_{i=1}^{n} |\, \langle e_i, x\rangle|^2 \\
&\leq \|x\|^2 - 2\sum_{i=1}^{n} \mathrm{Re}\ \langle e_i, x\rangle\ \langle x, e_i\rangle + \sum_{i=1}^{n} |\, \langle e_i, x\rangle|^2 \\
&\leq \|x\|^2 - 2\sum_{i=1}^{n} |\, \langle e_i, x\rangle|^2 + \sum_{i=1}^{n} |\, \langle e_i, x\rangle|^2 \\
&\leq \|x\|^2 - \sum_{i=1}^{n} |\, \langle e_i, x\rangle|^2
\end{aligned}
$$

no que implica

$$\sum_{i=1}^{n} |\, \langle e_i, x\rangle|^2 \leq \|x\|^2$$

Fazendo $n \to \infty$, a demonstração estará completa.

■

1.4 Ortonormalização de Gram-Schmidt

Teorema 1.15. *Sejam $x_1, x_2, x_3, ...$ finitos ou infinitos elementos pertencentes a um espaço unitário A. Seja também $L(x_1, ..., x_n)$ o espaço gerado pelos n primeiros elementos daquela sequência. Para todo $n \geq 1$, existe uma sequência $y_1, y_2, y_3, ...$ de elementos pertencentes a A tal que:*

1. *y_n é ortogonal aos elementos $y_1, y_2, y_3, ..., y_{n-1}$ (e portanto é ortogonal a qualquer elemento de $L(y_1, y_2, y_3, ..., y_{n-1})$)*
2. *$L(y_1, ..., y_n) = L(x_1, ..., x_n)$*
3. *Se existir uma outra sequência de elementos $w_1, w_2, w_3, \ldots$ de A que obedeça os itens 1 e 2 acima, então para todo n existe algum escalar c_n tal que*

$$w_n = c_n y_n$$

Demonstração. Definimos $y_1 = x_1$. Os outros elementos y_{r+1}, $r \geq 1$, serão definidos pela fórmula

$$y_{r+1} = x_{r+1} - \sum_{i=1}^{r} a_i y_i \tag{23}$$

onde os escalares a_i ainda são desconhecidos. No caso particular $r = 1$, temos

$$y_{1+1} = y_2 = x_2 - a_1 y_1$$

Então

$$\langle y_1, y_2 \rangle = \langle y_1, x_2 - a_1 y_1 \rangle = \langle y_1, x_2 \rangle - \langle y_1, a_1 y_1 \rangle = \langle x_1, x_2 \rangle - a_1 \langle x_1, x_1 \rangle$$

Queremos que $\langle y_1, y_2 \rangle = 0$, o que implica

$$\langle x_1, x_2 \rangle - a_1 \langle x_1, x_1 \rangle = 0 \implies a_1 = \frac{\langle x_1, x_2 \rangle}{\langle x_1, x_1 \rangle}$$

Suponha que os elementos y_i são ortogonais entre si. Seja y_j, $j \leq r$ um elemento intermediário de y_1 a y_r. Então o produto interno de y_j com y_{r+1} é dado por

$$\langle y_j, y_{r+1} \rangle = \langle y_j, x_{r+1} - \sum_{i=1}^{r} a_i y_i \rangle = \langle y_j, x_{r+1} \rangle - \langle y_j, \sum_{i=1}^{r} a_i y_i \rangle =$$
$$= \langle y_j, x_{r+1} \rangle - \sum_{i=1}^{r} \langle y_j, a_i y_i \rangle = \langle y_j, x_{r+1} \rangle - a_j \langle y_j, y_j \rangle$$

Como $j \neq r+1$, então $y_j \neq y_{r+1}$ e

$$\langle y_j, y_{r+1} \rangle = 0 \implies \langle y_j, x_{r+1} \rangle - a_j \langle y_j, y_j \rangle = 0$$

$$a_j = \frac{\langle y_j, x_{r+1} \rangle}{\langle y_j, y_j \rangle}$$

Se porventura $y_j = 0$, pelo corolário do axioma (13), y_j será ortogonal a todos os outros elementos. Então a escolha de a_j é arbitrária e definimos $a_j = 0$ para este caso.

Isso significa que existem escalares a_i tais que y_{r+1} seja ortogonal a todos os demais elementos e, portanto, também o é para o espaço gerado

$$L(y_1, ..., y_r)$$

provando assim o primeiro item do teorema, com $n = r + 1$.

Dado que $y_1 = x_1$, segue que $L(y_1) = L(x_1)$. Então suponha que

$$L(y_1, ..., y_r) = L(x_1, ..., x_r)$$

para algum $r \geq 1$. Sabemos que os primeiros r elementos y_i estão em $L(x_1, ..., x_r)$, e portanto também os estão em $L(x_1, ..., x_r, x_{r+1})$. Sabemos que, pela equação (23), y_{r+1} é uma combinação linear de x_{r+1} e dos elementos y_i, que por sua vez estão em $L(x_1, ..., x_r, x_{r+1})$. Segue que

$$L(y_1, ..., y_r, y_{r+1}) \subseteq L(x_1, ..., x_r, x_{r+1})$$

Similarmente, da equação (23) obtemos

$$x_{r+1} = y_{r+1} + \sum_{i=1}^{r} a_i y_i \tag{24}$$

O que significa que x_{r+1} é uma combinação linear dos $r+1$ elementos y_i e portanto

$$L(x_1, ..., x_r, x_{r+1}) \subseteq L(y_1, ..., y_r, y_{r+1})$$

Logo $L(x_1, ..., x_r, x_{r+1}) = L(y_1, ..., y_r, y_{r+1})$, provando o segundo item com $n = r + 1$.

Por indução, conseguimos encontrar n elementos y_i tais que obedeçam os dois primeiros itens.

Sejam y'_i outros elementos que satisfaçam os dois primeiros itens do teorema. Sabemos que $y'_1 = x_1 = 1 \cdot y_1$. Suponha que $y'_r = b_r y_r$ para algum $r \geq 1$. Como $L(y'_1, ..., y'_{r+1}) = L(x_1, ..., x_{r+1}) = L(y_1, ..., y_{r+1})$, segue que $y'_{r+1} \in L(y_1, ..., y_{r+1})$ e

$$y'_{r+1} = \sum_{i=1}^{r+1} c_i y_i = \sum_{i=1}^{r} (c_i y_i) + c_{r+1} y_{r+1} = z_r + c_{r+1} y_{r+1}$$

$$z_r = y'_{r+1} - c_{r+1} y_{r+1}$$

Pelo primeiro item do teorema, os elementos y_i, $1 \leq i \leq r$, são ortogonais a y_{r+1}, ou seja, $\langle z_r, y_{r+1} \rangle = 0$. Além disso, z_r é ortogonal a y'_{r+1}, visto que y'_{r+1} é ortogonal a $L(y'_1, ..., y'_r)$, logo, pelo segundo item, também o é para $z_r \in L(y_1, ..., y_r)$.

Calculando o produto interno de z_r com ele mesmo obtemos

$$\langle z_r, z_r \rangle = \langle z_r, y'_{r+1} - c_{r+1} y_{r+1} \rangle = \underbrace{\langle z_r, y'_{r+1} \rangle}_{=0} - \underbrace{\langle z_r, c_{r+1} y_{r+1} \rangle}_{=0} = 0$$

Pelo axioma (14), segue que $z_r = 0$ e portanto $y'_{r+1} = c_{r+1} y_{r+1}$. Logo o item 3 foi demonstrado pelo principio da indução.

■

Se $y_{r+1} = 0$, segue que

$$x_{r+1} = \sum_{i=1}^{r} a_i y_i$$

ou seja, x_{r+1} é uma combinação linear dos primeiros r elementos y_i, que por sua vez geram o mesmo espaço que os primeiros r elementos x_i. Isso implica que o conjunto $\{x_1, ..., x_{r+1}\}$ é dependente. Pela contrapositiva dessa

conclusão, temos que se $\{x_1, ..., x_{r+1}\}$ é independente, então todos os primeiros $r+1$ elementos y_i são não-nulos. Neste caso, obtemos uma sequência de n elementos ortogonais y_i não-nulos de um espaço unitário conhecendo n elementos x_i independentes entre si desse mesmo espaço. Isso pode ser feito usando essas fórmulas:

$$y_1 = x_1 \qquad y_{r+1} = x_{r+1} - \sum_{i=1}^{r} \frac{\langle y_i, x_{r+1}\rangle}{\langle y_i, y_i\rangle} \cdot y_i \quad \text{para } r = 1, 2, 3, ..., n-1 \tag{25}$$

Essas equações descrevem o método de Gram-Schmidt para encontrar um conjunto ortogonal de n elementos. Uma vez que o conjunto $B = \{x_1, ..., x_n\}$ é linearmente independente, então B é uma base de um espaço unitário A de dimensão n. Como o conjunto $C = \{y_1, ..., y_n\}$ é ortogonal, então, pelo teorema (1.11), C também é uma base de A. Logo o método de Gram-Schmidt também serve para encontrar bases ortogonais de qualquer espaço unitário de dimensão finita. Isso implica no teorema a seguir.

Teorema 1.16. *Todo espaço unitário que possui dimensão finita tem uma base ortonormal.*

Para encontrar uma base ortonormal, basta usar o método de Gram-Schmidt para obter uma base ortogonal e então dividir cada elemento pela sua respectiva norma.

1.4.1 Projeções

Seja A um espaço unitário de dimensão n e $C \subseteq A$ um subespaço unitário de A com dimensão $k \leq n$. Pelo teorema (1.16), C tem uma base ortonormal $B = \{e_1, ..., e_k\}$. Seja $b \in A$, $b \neq 0$. Se $k = n$, então b pode ser gerado pela base B. Se $k < n$, ou $b \in L(B)$ ou $b \notin L(B)$. No primeiro caso, b pode ser gerado por B. Mas se $b \notin L(B)$, nenhuma combinação linear dos elementos de B pode gerar b.

Entretanto, nesse caso é possível encontrar um elemento $w \in L(B)$ tal que a norma $\|b - w\|$ seja mínima. Seja e_{k+1} um novo elemento gerado pelo método de Gram-Schmidt em (23) com $x_{k+1} = b$. Isso significa que a nova base ortogonal $B' = \{e_1, ..., e_{k+1}\}$ gera b. Logo

$$b = c_{k+1}e_{k+1} + ... + c_1e_1 \tag{26}$$

Seja w o elemento $c_ke_k + ... + c_1e_1 \in L(B)$. Substituindo na equação (26) obtemos

$$b = c_{k+1}e_{k+1} + w \tag{27}$$

Então

$$
\begin{aligned}
\langle b, c_{k+1}e_{k+1}\rangle &= \langle c_{k+1}e_{k+1} + w, c_{k+1}e_{k+1}\rangle \\
c_{k+1}\langle b, e_{k+1}\rangle &= \langle c_{k+1}e_{k+1}, c_{k+1}e_{k+1}\rangle + \langle w, c_{k+1}e_{k+1}\rangle \\
c_{k+1}\langle b, e_{k+1}\rangle &= c_{k+1}\overline{c_{k+1}}\langle e_{k+1}, e_{k+1}\rangle + \langle w, c_{k+1}e_{k+1}\rangle
\end{aligned}
$$

Uma vez que a base B' é ortogonal, segue que

$$\langle w, c_{k+1}e_{k+1}\rangle = 0$$

Logo

$$
\begin{aligned}
c_{k+1}\langle b, e_{k+1}\rangle &= c_{k+1}\overline{c_{k+1}}\|e_{k+1}\|^2 \\
\langle b, e_{k+1}\rangle &= \overline{c_{k+1}}\|e_{k+1}\|^2 \\
\overline{\langle b, e_{k+1}\rangle} &= \overline{\overline{c_{k+1}}}\|e_{k+1}\|^2 \\
\frac{\langle e_{k+1}, b\rangle}{\|e_{k+1}\|^2} &= c_{k+1} \qquad (28)
\end{aligned}
$$

Conhecendo b e $c_{k+1}e_{k+1}$, então encontramos w:

$$w = b - c_{k+1}e_{k+1} \qquad (29)$$

Uma propriedade de w é que ele é único, pois se

$$w' = b - d_{k+1}e_{k+1}$$

então

$$
\begin{aligned}
w - w' &= b - c_{k+1}e_{k+1} - b + d_{k+1}e_{k+1} \\
w - w' &= (d_{k+1} - c_{k+1})e_{k+1}
\end{aligned}
$$

Mas $(w - w')$ não possui componente e_{k+1} já que $(w - w') \in L(B)$, o que implica que o escalar $d_{k+1} - c_{k+1}$ é nulo, ou seja, $w = w'$.

Lembremos ainda a equação (23):

$$e_{k+1} = b - \sum_{i=1}^{k} \langle e_i, b\rangle e_i$$

Então a equação para w fica

$$w = b - \left\langle b - \sum_{i=1}^{k} \langle e_i, b\rangle e_i, b \right\rangle \cdot \left(b - \sum_{i=1}^{k} \langle e_i, b\rangle e_i \right)$$

Note ainda que

$$b = w + \left\langle b - \sum_{i=1}^{k} \langle e_i, b \rangle \, e_i, b \right\rangle \cdot \left(b - \sum_{i=1}^{k} \langle e_i, b \rangle \, e_i \right)$$

Aplicando o produto interno de algum elemento e_m da base B com b e lembrando que todos os elementos de B estão normalizados, obtemos

$$\begin{aligned}
\langle e_m, b \rangle &= \langle e_m, w \rangle + \left\langle e_m, \left\langle b - \sum_{i=1}^{k} \langle e_i, b \rangle \, e_i, b \right\rangle \cdot \left(b - \sum_{i=1}^{k} \langle e_i, b \rangle \, e_i \right) \right\rangle \\
&= \langle e_m, w \rangle + \left\langle b - \sum_{i=1}^{k} \langle e_i, b \rangle \, e_i, b \right\rangle \cdot \left\langle e_m, \left(b - \sum_{i=1}^{k} \langle e_i, b \rangle \, e_i \right) \right\rangle \\
&= \langle e_m, w \rangle + \left\langle b - \sum_{i=1}^{k} \langle e_i, b \rangle \, e_i, b \right\rangle \cdot \left[\langle e_m, b \rangle - \left\langle e_m, \sum_{i=1}^{k} \langle e_i, b \rangle \, e_i \right\rangle \right] \\
&= \langle e_m, w \rangle + \left\langle b - \sum_{i=1}^{k} \langle e_i, b \rangle \, e_i, b \right\rangle \cdot \left[\langle e_m, b \rangle - \langle e_m, \langle e_m, b \rangle \, e_m \rangle \right] \\
&= \langle e_m, w \rangle + \left\langle b - \sum_{i=1}^{k} \langle e_i, b \rangle \, e_i, b \right\rangle \cdot \left[\langle e_m, b \rangle - \langle e_m, b \rangle \, \langle e_m, e_m \rangle \right] \\
&= \langle e_m, w \rangle + \left\langle b - \sum_{i=1}^{k} \langle e_i, b \rangle \, e_i, b \right\rangle \cdot \left[\langle e_m, b \rangle - \langle e_m, b \rangle \right] \\
&= \langle e_m, w \rangle
\end{aligned}$$

Uma vez que

$$w = \sum_{i=1}^{k} \langle e_i, w \rangle \, e_i$$

temos então

$$w = \sum_{i=1}^{k} \langle e_i, b \rangle \, e_i$$

Outra propriedade de w é ser o elemento de $L(B)$ tal que $\|b - w\|$ é mínimo. Para um $z \in L(B)$, note que

$$b - z = d_{k+1} e_{k+1} + u$$

para algum escalar d_{k+1} e algum elemento $u \in L(B)$. $d_{k+1} \neq 0$, pois $b - z \notin L(B)$ e $b - z \in L(B')$. A norma $\|b - z\|$ é minimizada quando $u = 0$, ou seja,

$$\begin{aligned}
b - z &= d_{k+1} e_{k+1} \\
z &= b - d_{k+1} e_{k+1} \qquad (30)
\end{aligned}$$

Comparando as equações (29) e (30) e sabendo que w é único, então concluímos que $z = w$ e $d_{k+1} = c_{k+1}$, e portanto w é o elemento de $L(B)$ que minimiza a norma $\|b - w\|$.

Definição 1.9. *Dado uma base ortonormal $B = \{e_1, ..., e_k\}$ de um subespaço euclidiano ou unitário C de A, $dim A = n$, $k \leq n$, e x um elemento de A, então o elemento*

$$\text{proj}_C\, x = \sum_{i=1}^{k} \langle e_i, x\rangle\, e_i \tag{31}$$

é a projeção de x no subespaço C.

Podemos supor também a seguinte expressão, com y_i sendo qualquer elemento não-nulo de A:

$$\begin{aligned} \frac{\langle y_i, x\rangle}{\langle y_i, y_i\rangle} y_i &= \frac{\langle y_i, x\rangle\, y_i}{\|y_i\| \cdot \|y_i\|} \\ &= \langle \frac{y_i}{\|y_i\|}, x\rangle\, \frac{y_i}{\|y_i\|} \\ &= \langle \hat{y}_i, x\rangle\, \hat{y}_i \end{aligned} \tag{32}$$

onde $\hat{y}$ é o elemento y normalizado. E então

$$\text{proj}_C\, x = \sum_{i=1}^{k} \langle \hat{y}_i, x\rangle\, \hat{y}_i = \sum_{i=1}^{k} \frac{\langle y_i, x\rangle}{\langle y_i, y_i\rangle} y_i \tag{33}$$

onde $C = L(\{y_1, \ldots, y_k\})$.

Comparando com o método de Gram-Schmidt, o próximo $r + 1$ elemento ortogonal é obtido pela diferença entre x_{r+1} e a projeção de x_{r+1} em $C = L(\{y_1, \ldots, y_r\})$. Então podemos reescrever a equação (25) como

$$y_{r+1} = x_{r+1} - \text{proj}_C(x_{r+1}) \tag{34}$$

1.4.2 Exemplos de polinômios ortogonais

Exemplo 1.4 (Polinômios de Legendre)**.** *Seja A_n o espaço euclidiano dos polinômios de grau menor ou igual a n, e o produto interno entre dois elementos f e g de A_n seja dado por*

$$\langle f, g\rangle = \int_{-1}^{1} f(x) g(x)\, \mathrm{d}x \tag{35}$$

Uma base que pode ser utilizada para o método de Gram-Schmidt é a base canônica $\xi = \{a_i,\ 0 \leq i \leq\ n\}$, onde $a_i = x^i$. Sabemos que $y_0 = a_0 = x^0 = 1$. Pela equação (25):

$$y_1 = a_1 - \frac{\langle y_0, a_1 \rangle}{\langle y_0, y_0 \rangle} y_0 = x - \frac{\langle 1, x \rangle}{\langle 1, 1 \rangle} \cdot 1$$

$$\langle 1, x \rangle = \int_{-1}^{1} x \, \mathrm{d}x = 0$$

$$y_1 = x$$

$$y_2 = a_2 - \frac{\langle y_1, a_2 \rangle}{\langle y_1, y_1 \rangle} y_1 - \frac{\langle y_0, a_2 \rangle}{\langle y_0, y_0 \rangle} y_0 = x^2 - \frac{\langle x, x^2 \rangle}{\langle x, x \rangle} x - \frac{\langle 1, x^2 \rangle}{\langle 1, 1 \rangle} 1$$

$$\langle x, x^2 \rangle = \int_{-1}^{1} x^3 \, \mathrm{d}x = 0$$

$$\langle 1, x^2 \rangle = \int_{-1}^{1} x^2 \, \mathrm{d}x = \left. \frac{x^3}{3} \right|_{-1}^{1} = \frac{2}{3}$$

$$\langle 1, 1 \rangle = \int_{-1}^{1} \mathrm{d}x = 2$$

$$y_2 = x^2 - \frac{1}{3}$$

$$\begin{aligned} y_3 &= a_3 - \frac{\langle y_2, a_3 \rangle}{\langle y_2, y_2 \rangle} y_2 - \frac{\langle y_1, a_3 \rangle}{\langle y_1, y_1 \rangle} y_1 - \frac{\langle y_0, a_3 \rangle}{\langle y_0, y_0 \rangle} y_0 \\ &= x^3 - \frac{\langle x^2 - \frac{1}{3}, x^3 \rangle}{\langle x^2 - \frac{1}{3}, x^2 - \frac{1}{3} \rangle} \left(x^2 - \frac{1}{3} \right) - \frac{\langle x, x^3 \rangle}{\langle x, x \rangle} x - \frac{\langle 1, x^3 \rangle}{\langle 1, 1 \rangle} 1 \end{aligned}$$

$$\left\langle x^2 - \frac{1}{3}, x^3 \right\rangle = \int_{-1}^{1} x^5 - \frac{x^3}{3} \, \mathrm{d}x = 0 \textit{ (função ímpar em intervalo simétrico)}$$

$$\langle x, x^3 \rangle = \int_{-1}^{1} x^4 \, \mathrm{d}x = \left. \frac{x^5}{5} \right|_{-1}^{1} = \frac{2}{5}$$

$$\langle x, x \rangle = \int_{-1}^{1} x^2 \, \mathrm{d}x = \left. \frac{x^3}{3} \right|_{-1}^{1} = \frac{2}{3}$$

$$\langle 1, x^3 \rangle = \int_{-1}^{1} x^3 \, \mathrm{d}x = 0$$

$$y_3 = x^3 - \frac{3}{5} x$$

e assim sucessivamente.

Uma vez que pelo axioma (13) um múltiplo de y_m também é ortogonal a todos os outros elementos y_i, então temos a faculdade de mudar arbitrariamente a norma (desde que $\|y_i\| \neq 0$) desses conforme a conveniência do problema tratado.

No caso dos polinômios de Legendre, há uma conveniência de fazer $P_i(1) = 1$ para todos os elementos. Para fazer isso, basta multiplicar cada polinômio y_i por $\dfrac{1}{y_i(1)}$.

Visto que

$$y_0(1) = 1 \quad y_1(1) = 1 \quad y_2(1) = \frac{2}{3} \quad y_3(1) = \frac{2}{5}$$

Então os primeiros 4 polinômios de Legendre na conveniência discutida são:

$$P_0(x) = 1 \quad P_1(x) = x \quad P_2(x) = \frac{3}{2}x^2 - \frac{1}{2} \quad P_3(x) = \frac{5}{2}x^3 - \frac{3}{2}x$$

Exemplo 1.5 (Polinômios de Hermite)**.** *Seja A_n o espaço euclidiano do exemplo* (1.4) *e o produto interno entre dois elementos f e g de A_n seja dado por*

$$\langle f, g \rangle = \int_{-\infty}^{\infty} e^{-x^2} f(x)g(x)\, \mathrm{d}x \tag{36}$$

Usaremos a mesma base ξ utilizada no exemplo (1.4) *para gerar os polinômios ortogonais.*

Primeiramente sabemos que

$$y_0 = a_0 = x^0 = 1$$

$$\int_{-\infty}^{\infty} x^n e^{-x^2}\, \mathrm{d}x = \begin{cases} 0 & \textit{se } n \textit{ é ímpar} \\ \dfrac{n!\sqrt{\pi}}{2^n \left(\dfrac{n}{2}\right)!} & \textit{se } n \textit{ é par} \end{cases} \tag{37}$$

Então:

$$y_1 = a_1 - \frac{\langle y_0, a_1 \rangle}{\langle y_0, y_0 \rangle} \cdot y_0 = x - \frac{\int_{-\infty}^{\infty} x^1 e^{-x^2}\, \mathrm{d}x}{\langle y_0, y_0 \rangle} \cdot 1 = x - \frac{0}{\langle y_0, y_0 \rangle} = x$$

$$y_2 = a_2 - \frac{\langle y_0, a_2\rangle}{\langle y_0, y_0\rangle} \cdot y_0 - \frac{\langle y_1, a_2\rangle}{\langle y_1, y_1\rangle} \cdot y_1$$

$$\langle y_0, a_2\rangle = \int_{-\infty}^{\infty} e^{-x^2} x^2 \, \mathrm{d}x = \frac{\sqrt{\pi}}{2}$$

$$\langle y_0, y_0\rangle = \int_{-\infty}^{\infty} e^{-x^2} \, \mathrm{d}x = \sqrt{\pi}$$

$$\langle y_1, a_2\rangle = \int_{-\infty}^{\infty} x^3 e^{-x^2} \, \mathrm{d}x = 0$$

$$y_2 = x^2 - \frac{1}{2}$$

$$y_3 = a_3 - \frac{\langle y_0, a_3\rangle}{\langle y_0, y_0\rangle} \cdot y_0 - \frac{\langle y_1, a_3\rangle}{\langle y_1, y_1\rangle} \cdot y_1 - \frac{\langle y_2, a_3\rangle}{\langle y_2, y_2\rangle} \cdot y_2$$

$$\langle y_0, a_3\rangle = \langle 1, x^3\rangle = \int_{-\infty}^{\infty} x^3 e^{-x^2} \, \mathrm{d}x = 0$$

$$\langle y_1, a_3\rangle = \langle x, x^3\rangle = \int_{-\infty}^{\infty} x^4 e^{-x^2} \, \mathrm{d}x = \frac{4!\sqrt{\pi}}{2^4 \left(\frac{4}{2}\right)} = \frac{3\sqrt{\pi}}{4}$$

$$\langle y_1, y_1\rangle = \langle x, x\rangle = \int_{-\infty}^{\infty} x^2 e^{-x^2} \, \mathrm{d}x = \frac{2!\sqrt{\pi}}{2^2 \left(\frac{2}{2}\right)} = \frac{\sqrt{\pi}}{2}$$

$$\langle y_2, a_3\rangle = \langle x^2 - \frac{1}{2}, x^3\rangle = \int_{-\infty}^{\infty} x^5 e^{-x^2} \, \mathrm{d}x - \frac{1}{2}\int_{-\infty}^{\infty} x^3 e^{-x^2} \, \mathrm{d}x = 0$$

$$\begin{aligned} y_3 &= x^3 - \frac{\frac{3\sqrt{\pi}}{4}}{\frac{\sqrt{\pi}}{2}} x \\ &= x^3 - \frac{3}{2} x \end{aligned}$$

E assim sucessivamente.

Em outras aplicações, os polinômios de Hermite se encontram ortogonalizados em relação à função peso $e^{-\frac{x^2}{2}}$ ao invés de e^{-x^2}. Fazendo $x = \dfrac{x'}{\sqrt{2}}$ na equação (37), obtemos

$$\int_{-\infty}^{\infty} x^n e^{-x^2} \, \mathrm{d}x = \frac{1}{\sqrt{2}} \int_{-\infty}^{\infty} \left(\frac{x'}{\sqrt{2}}\right)^n e^{-\left(\frac{x'}{\sqrt{2}}\right)^2} \, \mathrm{d}x' = \left(\frac{1}{\sqrt{2}}\right)^{n+1} \int_{-\infty}^{\infty} x^n e^{-\frac{x^2}{2}} \, \mathrm{d}x$$

E então, se quisermos alterar a função peso do produto interno da equação (36) para $e^{-\frac{x^2}{2}}$, podemos usar as projeções da base canônica já encontradas anteriormente e usar a seguinte fórmula:

$$\int_{-\infty}^{\infty} x^n e^{-\frac{x^2}{2}}\,\mathrm{d}x = \sqrt{2^{n+1}} \int_{-\infty}^{\infty} x^n e^{-x^2}\,\mathrm{d}x$$

Isso significa que se os polinômios com o peso e^{-x^2} já estiverem ortogonalizados, então para encontrar os polinômios ortogonais com o peso $e^{-\frac{x^2}{2}}$ basta multiplicar cada projeção $\frac{\langle y_i, a_j\rangle}{\langle y_i, y_i\rangle} y_i$ por $\sqrt{\frac{2^{i+j+1}}{2^{2i+1}}} = \sqrt{2^{j-i}}$

Exemplo 1.6 (Polinômios de Laguerre)**.** *Seja A_n o espaço euclidiano dos exemplos* (1.4) *e* (1.5) *e o produto interno entre dois elementos ϕ e θ de A_n seja dado por*

$$\langle \phi, \theta\rangle = \int_0^{\infty} e^{-x}\phi(x)\theta(x)\,\mathrm{d}x \tag{38}$$

Os dois primeiros polinômios ortonormais gerados pela base canônica $\xi = \{1, x, x^2, \ldots, x^n\}$ são

$$y_0 = 1$$

$$y_1 = x - \frac{\langle 1, x\rangle}{\langle 1, 1\rangle} 1$$
$$\langle 1, x\rangle = \int_0^{\infty} x e^{-x}\,\mathrm{d}x$$
$$\langle 1, 1\rangle = \int_0^{\infty} e^{-x}\,\mathrm{d}x$$

Devemos observar que

$$\int_0^{\infty} e^{-x}\,\mathrm{d}x = -e^{-x}\Big|_0^{\infty} = 1 \tag{39}$$

Seja t uma constante positiva em relação a x. Fazendo $x = tx'$, obtemos $\mathrm{d}x = t\,\mathrm{d}x'$, $x \longrightarrow 0 \implies x' \longrightarrow 0$, $x \longrightarrow \infty \implies x' \longrightarrow \infty$ *e*

$$\int_0^{\infty} e^{-x}\,\mathrm{d}x = \int_0^{\infty} t e^{-tx'}\,\mathrm{d}x' = 1$$
$$\int_0^{\infty} e^{-tx'}\,\mathrm{d}x' = \frac{1}{t} \tag{40}$$

Aplicando o operador $\frac{\mathrm{d}}{\mathrm{d}t}$ *a ambos os lados da equação* (40) *e mudando a variável muda* x' *para* x*, obtemos*

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\int_0^\infty e^{-tx}\,\mathrm{d}x &= \frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{t} \\
\int_0^\infty \frac{\partial}{\partial t}e^{-tx}\,\mathrm{d}x &= -\frac{1}{t^2} \\
\int_0^\infty -xe^{-tx'}\,\mathrm{d}x &= -\frac{1}{t^2} \\
\int_0^\infty xe^{-tx}\,\mathrm{d}x &= \frac{1}{t^2} \qquad (41)
\end{aligned}$$

Aplicando o operador novamente na equação (41)*, temos*

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\int_0^\infty xe^{-tx}\,\mathrm{d}x &= \frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{t^2} \\
\int_0^\infty x^2e^{-tx}\,\mathrm{d}x &= \frac{1\cdot 2}{t^3} \qquad (42)
\end{aligned}$$

Repetindo o mesmo procedimento $n-2$ *vezes na equação* (42) *obtemos*

$$\int_0^\infty x^ne^{-tx}\,\mathrm{d}x = \frac{n!}{t^{n+1}}$$

Fazendo $t = 1$*, chegamos na expressão*

$$\int_0^\infty x^ne^{-x}\,\mathrm{d}x = n! \qquad (43)$$

E então podemos calcular os seguintes produtos internos

$$\begin{aligned}
\langle 1, x\rangle &= \int_0^\infty xe^{-x}\,\mathrm{d}x = 1! = 1 \\
\langle 1, 1\rangle &= \int_0^\infty e^{-x}\,\mathrm{d}x = 0! = 1
\end{aligned}$$

Dessa forma, podemos calcular os outros polinômios:

$$y_1 = x - 1$$

$$y_2 = x^2 - \frac{\langle x-1, x^2\rangle}{\langle x-1, x-1\rangle}(x-1) - \frac{\langle 1, x^2\rangle}{\langle 1, 1\rangle}1$$

$$
\begin{aligned}
\langle x-1, x^2\rangle &= \int_0^\infty (x-1)x^2 e^{-x}\,\mathrm{d}x \\
&= \int_0^\infty x^3 e^{-x}\,\mathrm{d}x - \int_0^\infty x^2 e^{-x}\,\mathrm{d}x \\
&= 3! - 2! \\
&= 4
\end{aligned}
$$

$$
\begin{aligned}
\langle x-1, x-1\rangle &= \langle x, x\rangle - \langle x, 1\rangle - \langle 1, x\rangle + \langle 1, 1\rangle \\
&= \int_0^\infty x^2 e^{-x}\,\mathrm{d}x - 2\int_0^\infty x e^{-x}\,\mathrm{d}x + \int_0^\infty e^{-x}\,\mathrm{d}x \\
&= 2! - 2\cdot(1!) + 0! \\
&= 1
\end{aligned}
$$

$$
\begin{aligned}
\langle 1, x^2\rangle &= 2! = 2 \\
\langle 1, 1\rangle &= 1! = 1
\end{aligned}
$$

$$
\begin{aligned}
y_2 &= x^2 - 4(x-1) - 2 \\
&= x^2 - 4x + 2
\end{aligned}
$$

$$
\begin{aligned}
y_3 = x^3 - \frac{\langle x^2-4x+2, x^3\rangle}{\langle x^2-4x+2, x^2-4x+2\rangle}(x^2-4x+2) \\
- \frac{\langle x-1, x^3\rangle}{\langle x-1, x-1\rangle}(x-1) - \frac{\langle 1, x^3\rangle}{\langle 1, 1\rangle}1
\end{aligned}
$$

$$
\langle x^2-4x+2, x^3\rangle = 5! - 4\cdot(4!) + 2\cdot(3!) = 36
$$

$$
\begin{aligned}
\langle x^2-4x+2, x^2-4x+2\rangle &= 4! - 8*(3!) + 20*(2!) - 16*(1!) + 4*(0!) \\
&= 4
\end{aligned}
$$

$$
\begin{aligned}
\langle x-1, x^3\rangle &= 4! - 3! = 18 \\
\langle 1, x^3\rangle &= 3! = 6
\end{aligned}
$$

$$
\begin{aligned}
y_3 &= x^3 - \frac{36}{4}(x^2 - 4x + 2) - 18 \cdot (x-1) - 6 \cdot 1 \\
&= x^3 - 9x^2 + 36x - 18 - 18x + 18 - 6 \\
&= x^3 - 9x^2 + 18x - 6
\end{aligned}
$$

E assim sucessivamente.

Em certas finalidades, os polinômios de Laguerre precisam obedecer a propriedade $y_i(0) = 1, \ \forall i \geq 0$. Fazendo $x = 0$ nos polinômios encontrados acima, obtemos

$$
\begin{aligned}
y_0(0) &= 1 \\
y_1(0) &= -1 \\
y_2(0) &= 2 \\
y_3(0) &= -6 \\
(\ldots)
\end{aligned}
$$

Note que $y_n(0) = (-1)^n n!$ para os 4 primeiros polinômios. Embora ainda não seja possível demonstrar essa relação para todos os infinitos polinômios descobertos acima pelo método de Gram-Schmidt, ainda é possível obter finitos polinômios naquela conveniência fazendo

$$P_n(x) = \frac{y_n(x)}{y_n(0)}$$

1.5 Transformações Lineares

Definição 1.10. *Seja $T : V \mapsto W$ uma função que leva os elementos de um espaço vetorial V para elementos do espaço vetorial W, ambos sobre o mesmo corpo K. Se T obedecer as propriedades da linearidade,*

1. $T(x+y) = T(x) + T(y), \quad \forall x, y \in V$
2. $T(cx) = cT(x), \quad \forall x \in V; \forall c \in K$

então T é chamado de transformação linear de V para W.

As duas propriedades podem ser combinadas pela equação (44)

$$T(cx + dy) = cT(x) + d(y), \quad \forall x; y \in V, \forall c, d \in K \tag{44}$$

Suponha que para um certo $n \geq 1$ a seguinte equação seja válida:

$$T\left(\sum_{i=0}^{n} d_i x_i\right) = \sum_{i=0}^{n} d_i T(x_i) \quad \forall x_i \in V; \forall d_i \in K \tag{45}$$

Então

$$T\left(\left(\sum_{i=0}^{n} d_i x_i\right) + d_{n+1}x_{n+1}\right) = T\left(\sum_{i=0}^{n} d_i x_i\right) + d_{n+1}T(x_{n+1}) = \sum_{i=0}^{n+1} d_i T(x_i)$$

Por indução, a propriedade da equação (44) pode ser estendida para um número arbitrário de elementos e escalares dados pela equação (45).

Exemplo 1.7 (Transformação identidade). *Seja V um espaço vetorial qualquer. A função $I : V \mapsto V, \quad I(x) = x, \quad \forall x \in V$ é uma transformação linear, pois*

- $I(x_0 + x_1) = x_0 + x_1 = I(x_0) + I(x_1)$
- $I(cx) = cx = cI(x),$ *para todo escalar c*

Exemplo 1.8 (Transformação nula). *Seja V um espaço vetorial qualquer. A função $T : V \mapsto V, \quad T(x) = 0, \quad \forall x \in V$ é uma transformação linear, pois*

- $T(x_0 + x_1) = 0 = 0 + 0 = T(x_0) + T(x_1)$
- $T(cx) = 0 = c0 = cT(x),$ *para todo escalar c*

Exemplo 1.9 (O operador derivada). *Seja V o espaço vetorial das funções deriváveis em um intervalo I de uma variável. O operador D que leva uma função f de V a sua derivada f' é uma transformação linear, visto que a derivada também possui as propriedades da linearidade.*

Exemplo 1.10 (O operador integral). *Seja V o espaço vetorial das funções integráveis em um intervalo $I = [a, b]$ de uma variável. O operador T que leva uma função $f \in V$ a F da seguinte forma*

$$F(x) = \int_a^x f(x')\,\mathrm{d}x' \qquad a \leq x \leq b$$

é uma transformação linear, já que a integral possui propriedades da linearidade.

1.5.1 Núcleo e imagem

Teorema 1.17. *Seja T uma transformação linear de V para W. Então as duas propriedades seguintes são válidas:*

1. *O conjunto imagem $T(V)$ é um subespaço vetorial de W.*
2. *A aplicação de T no elemento nulo de V leva ao elemento nulo de W.*

Demonstração. Sejam $x, y \in V$, a e b escalares. Então

$$aT(x) + bT(y) = T(ax + by) \in T(V)$$

o que prova o primeiro item.

Seja v_0 o elemento zero de V e w_0 o elemento zero de W. Então

$$\begin{aligned} T(v_0) = T(v_0 + v_0) &= T(v_0) + T(v_0) \\ w_0 &= T(v_0) \end{aligned}$$

■

Nesta e na próxima seção, pressuponha que T seja uma transformação linear de V para W.

Definição 1.11. *Designa-se núcleo de T o conjunto N de todos os elementos de V tais que $T(N) = \{0\}$.*

Teorema 1.18. *O núcleo de T é um subespaço vetorial de V.*

Demonstração. Seja $x, y \in N$, a e b escalares. Então

$$aT(x) + bT(y) = T(ax + by)$$

Mas

$$\begin{aligned} aT(x) &= a \cdot 0 = 0 \\ bT(y) &= b \cdot 0 = 0 \end{aligned}$$

Logo

$$T(ax + by) = 0 + 0 = 0$$

o que implica $ax + by \in N$.

■

Denominamos o Núcleo como $N(T)$.

Exemplo 1.11. *O núcleo do operador derivada é composto por todas as funções constantes no intervalo I.*

Exemplo 1.12. *O núcleo do operador integral é composto unicamente pela função nula.*

1.5.2 Nulidade e ordem

Se a dimensão do núcleo da transformação linear T for finita, então denominamos $\dim N(T)$ a nulidade de T. Se a dimensão da imagem $T(V)$ for finita, denominamos $\dim T(V)$ a ordem de T.

Teorema 1.19. *Se V tem dimensão finita, então $T(V)$ tem dimensão finita e*

$$\dim N(T) + \dim T(V) = \dim V \tag{46}$$

Demonstração. Seja o subconjunto $C = \{e_0, e_1, \ldots, e_k\}$ uma base para o núcleo $N(T)$ associado a V, de dimensão n. Uma vez que o núcleo é um subespaço de V, então $\dim N(T) \leq \dim V$. Além disso, pelo teorema (1.7), C faz parte de alguma outra base de V, $B = \{e_0, e_1, \ldots, e_k, e_{k+1}, \ldots, e_{k+r}\}$. Pelo teorema (1.7), $k+r = n$. Seja y um elemento qualquer de $T(V)$. Então existe algum $x \in V$ tal que $T(x) = y$. Como $x \in L(B)$, temos que

$$x = \sum_{i=0}^{k} c_i e_i + \sum_{i=k+1}^{r+k} c_i e_i$$

Pela linearidade de T:

$$y = T(x) = T\left(\sum_{i=0}^{k} c_i e_i\right) + T\left(\sum_{i=k+1}^{r+k} c_i e_i\right) = \sum_{i=0}^{k} c_i T(e_i) + \sum_{i=k+1}^{r+k} c_i T(e_i)$$

Uma vez que $T(e_i) = 0$ se $0 \leq i \leq k$, então $\sum_{i=0}^{k} c_i T(e_i) = 0$ e

$$y = \sum_{i=k+1}^{r+k} c_i T(e_i)$$

o que implica que qualquer elemento da imagem pode ser gerado pela transformação das combinações lineares do subconjunto $A = \{e_{k+1}, \ldots, e_{k+r}\}$, de dimensão r. Em outras palavras, $D = \{T(e_{k+1}), \ldots, T(e_{k+r})\}$ gera $T(V)$.

Para mostrarmos que $T(V)$ tem dimensão finita, é preciso encontrar uma base finita de $T(V)$. Usando o subconjunto A na definição (1.2) obtemos

$$\sum_{i=k+1}^{r+k} d_i T(e_i) = 0$$

$$T\left(\sum_{i=k+1}^{r+k} d_i e_i\right) = 0$$

Mas nenhum elemento gerado por $A = \{e_{k+1}, \ldots, e_{k+r}\}$ aplicado por T pode ser igual ao elemento nulo (exceto ele próprio), visto que nenhum destes

pertencem ao núcleo. Sabendo do segundo item do teorema (1.17), a única alternativa restante é que

$$\sum_{i=k+1}^{r+k} d_i e_i = 0$$

Mas A é um conjunto linearmente independente, uma vez que, pelo teorema (1.7), é um subconjunto de um conjunto independente B. Logo $d_i = 0\ \forall i,\ k+1 \leq i \leq k+r$. Consequentemente $D = \{T(e_{k+1}), \ldots, T(e_{k+r})\}$ é linearmente independente. Isso significa que pela definição (1.3), D é uma base finita de $T(V)$, e portanto $T(V)$ tem dimensão finita. Sabendo que D tem dimensão r e que $k + r = n$, segue que

$$\dim N(T) + \dim T(V) = \dim V$$

■

1.5.3 Operações algébricas com transformações lineares

Definição 1.12. *Sejam $S : V \mapsto W$ e $T : V \mapsto W$ duas transformações lineares de mesmo domínio V. Seja também c um escalar de W. Definimos $T + S$ e cT respectivamente como a soma e o produto por escalar da seguinte forma:*

$$(T + S)(x) = T(x) + S(x) \qquad cT = cT(x) \qquad \forall x \in V \tag{47}$$

Se o escalar c for de um corpo comum a ambos V e W, então denominamos $\mathcal{L}(V, W)$ o conjunto de todas as transformações lineares de V em W.

Ambas as transformações $T + S$ e cT pertencem a $\mathcal{L}(V, W)$, o que significa que as operações de soma e produto são fechadas. Além disso, possui uma transformação $0 : V \mapsto W,\ 0(x) = 0 \in W$ tal que $T(x) + 0(x) = T(x)$ para toda transformação em $\mathcal{L}(V, W)$, e ainda uma transformação simétrica $(-1)T : V \mapsto W,\ (-1)T(x) = 0 - T(x)$ tal que $T(x) + (-1)T(x) = 0(x)$ para cada $T(x) \in \mathcal{L}(V, W)$. Os demais axiomas do espaço vetorial são satisfeitos, uma vez que V e W são espaços vetoriais. Isso culmina no teorema a seguir.

Teorema 1.20. *O conjunto $\mathcal{L}(V, W)$ de todas as transformações lineares de V em W com as operações de soma e multiplicação dadas pelas equações* (47) *é um espaço vetorial.*

Definição 1.13. *Sejam U, V e W conjuntos e as funções $T : U \mapsto V$ e $S : V \mapsto W$. A composição ST é definida por*

$$ST : U \mapsto W,\ ST(x) = S(T(x)),\ \forall x \in U$$

Teorema 1.21. *Dadas as funções $T : U \mapsto V,\ S : V \mapsto W$ e $R : W \mapsto X$ é verdade que*

$$R(ST) = (RS)T$$

Demonstração. Ambas as transformações $R(ST)$ e $(RS)T$ possuem domínio em U e imagem em X. Ainda

$$\begin{aligned} R(ST) &= R[S(T(x))] = R(S(T(x))) \qquad \forall x \in U \\ (RS)T &= (RS)(T(x)) = R(S(T(x))) \qquad \forall x \in U \end{aligned}$$

Logo $R(ST) = (RS)T$. ■

Definição 1.14. *Seja $T : V \mapsto V$ uma transformação linear. Definimos as potências não-negativas de T da seguinte forma:*

$$T^n(x) = \begin{cases} I(x) & , \text{ se } n = 0 \\ T(T^{n-1}(x)) & , \text{ se } n > 0 \end{cases} \qquad \forall x \in V$$

onde $I(x)$ é a transformação identidade do exemplo (1.7) *e n é um inteiro não-negativo.*

Observe que

$$T^mT^n = T^m(T^n(x)) = T^m(x), \text{ se } n = 0$$

$$T^mT^n = T^m(T(T^{n-1}(x))) = T^{m+1}(T^{n-1}(x)) = T^{m+1}T^{n-1}$$

Dado que n é um inteiro não-negativo, existe algum inteiro não-negativo r tal que $n - r = 0$. Repetindo o processo acima $r - 1$ vezes obtemos

$$T^{m+r}T^{n-r} = T^{m+r}T^0 = T^{m+r}$$

Mas $n - r = 0 \implies r = n$, logo

$$T^{m+r}T^{n-r} = T^{m+n} \implies T^mT^n = T^{m+n}$$

Teorema 1.22. *Dado os espaços vetoriais U, V e W cujos escalares são de um corpo comum K, se as funções $T : U \mapsto V$ e $S : V \mapsto W$ forem transformações lineares, então a função $ST : U \mapsto V$ é também uma transformação linear.*

Demonstração. Dado $x, y \in U$ e $c \in K$, temos

$$ST(x+y) = S(T(x+y)) = S(T(x)+T(y)) = S(T(x))+S(T(y)) = ST(x)+ST(y)$$

$$ST(cx) = S(T(cx)) = S(cT(x)) = cS(T(x)) = cST(x)$$

■

Teorema 1.23. *Dado os espaços vetoriais U, V e W cujos escalares são de um corpo comum K, as transformações lineares $S, T \in \mathcal{L}(V, W)$ e $c \in K$. Então é verdade que*

1. *Para qualquer função R cuja imagem é subconjunto de V, tem-se*

$$(S+T)R = SR + TR \quad e \quad (cS)R = c(SR)$$

2. *Para qualquer transformação linear $R : W \mapsto U$, tem-se*

$$R(S+T) = RS + RT \quad e \quad R(cS) = c(RS)$$

Demonstração. Para o primeiro item, suponha que y pertence ao domínio de R. Então

$$(S+T)R = (S+T)(R(y)) = S(R(y)) + T(R(y)) = SR + TR$$
$$(cS)R = (cS)(R(y)) = cS(R(y)) = c(SR)$$

Para o segundo item, suponha que $x \in V$. Então

$$R(S+T) = R(S(x)+T(x)) = R(S(x)) + R(T(x))^{**} = RS + RT$$

$$R(cS) = R(cS(x)) = cR(S(x))^{**} = c(RS)$$

*** R é uma transformação linear.* ■

1.5.4 Inversas

Definição 1.15. *Sejam V e W conjuntos e a função $T : V \mapsto W$. Uma função $L : T(V) \mapsto V$ tal que $L(T(x)) = x$, $\forall x \in V$ é chamada de inversa esquerda de T. Uma função $R : T(V) \mapsto V$ tal que $T(R(y)) = y$, $\forall y \in T(V)$ é dita inversa direita de T.*

Teorema 1.24. *Seja a função $T : V \mapsto W$. Se T tem uma inversa esquerda, então ela é única e também é uma inversa direita de T.*

Demonstração. Suponha que $L_1 : T(V) \mapsto V$ e $L_2 : T(V) \mapsto V$ são duas inversas esquerdas de T. Então

$$L_1(T(x)) = x \quad \text{e} \quad L_2(T(x)) = x \quad \forall x \in V$$

Subtraindo as duas funções obtemos

$$\begin{aligned} L_1(T(x)) - L_2(T(x)) &= x - x \\ &= 0 \\ L_1(T(x)) &= L_2(T(x)) \quad \forall x \in V \end{aligned}$$

Dado que a relação e os domínios de L_1 e L_2 são iguais, então podemos afirmar que a inversa esquerda L de T, se existir, é única.

Além disso, supondo que $T(x) = y \quad$ (o que também implica $L(y) = x$) e aplicando T no resultante de L obtemos

$$T(L(T(x))) = T(L(y)) = T(x) \quad \forall x \in V$$

Reescrevendo a equação anterior,

$$T(L(y)) = y \quad \forall y \in T(V)$$

o que significa que a inversa esquerda de T também é uma inversa direita de T. ■

Teorema 1.25. *Uma função $T : V \mapsto W$ só possui uma inversa esquerda se e somente se T for injetora, ou seja*

$$T(x_1) = T(x_2) \implies x_1 = x_2$$

Se T for injetora, chamamos-a de biunívoca em V.

Demonstração. Se T possui uma inversa esquerda L, então dado $x_1, x_2 \in V$ temos que

$$T(x_1) = T(x_2) \implies L(T(x_1)) = L(T(x_2)) \implies x_1 = x_2$$

Por outro lado, suponha que $T(x_1) = T(x_2) \implies x_1 = x_2$. Seja a função $L : T(V) \mapsto V,\ L(y) = x$ tal que $T(x) = y$. Se $T(x_1) = T(x_2)$, então $L(T(x_1)) = L(T(x_2))$. A função só pode existir se $x_1 = x_2$, o que é assegurado, uma vez que $T(x_1) = T(x_2) \implies x_1 = x_2$. Visto que $L(T(x)) = x,\ \forall x \in V$, então L é uma inversa esquerda de T, o que completa a prova. ■

Definição 1.16. *Seja a função $T : V \mapsto W$ biunívoca em V. Então a função é dita invertível e a sua única inversa esquerda representa-se por T^{-1}.*

Teorema 1.26. *Seja a transformação linear $T : V \mapsto W$, $T \in \mathcal{L}(V, W)$. Então são afirmações equivalentes:*

1. *T é biunívoca em V.*
2. *T é invertível e T^{-1} é uma transformação linear.*
3. *$N(T) = \{0\}$*

Demonstração. Provando (*1*) $\implies$ (*2*). Se a transformação linear T é biunívoca em V, então pelo teorema (1.25) T tem uma inversa esquerda T^{-1}. Além disso,

$$\begin{aligned} T^{-1}(T(x_1) + T(x_2)) &= T^{-1}(T(x_1 + x_2)) \\ &= x_1 + x_2 \\ &= T^{-1}(T(x_1)) + T^{-1}(T(x_2)) \quad \forall x_1, x_2 \in V \end{aligned}$$

$$T^{-1}(cT(x)) = T^{-1}(T(cx)) = cx = cT^{-1}(T(x)) \quad \text{para todo escalar } c \text{ e } \forall x \in V$$

Logo se T é linear e biunívoca, então a sua inversa esquerda T^{-1} também é linear.

Provando (*2*) $\implies$ (*3*). Suponha que T seja invertível e que T^{-1} é linear. Suponha também que para algum $x \in V$, $T(x) = 0$. Aplicando a inversa em $T(x)$ resulta

$$\begin{aligned} T^{-1}(T(x)) &= T^{-1}(0) \\ x &= 0 \end{aligned}$$

A condição $T^{-1}(0) = 0$ é garantida pelo teorema (1.17), visto que T^{-1} é linear. Consequentemente, $T(x) = 0 \implies x = 0$, ou seja, $N(T) = \{0\}$.

Provando (*3*) $\implies$ (*1*). Agora se uma transformação linear $T : V \mapsto W$, $S \in \mathcal{L}(V, W)$ possui a propriedade $N(T) = \{0\}$, ou seja, $T(x) = 0 \implies x = 0$ para qualquer $x \in V$, então

$$\begin{aligned} T(x_1 - x_2) = 0 &\implies x_1 - x_2 = 0 \\ T(x_1) - T(x_2) = 0 &\implies x_1 = x_2 \\ T(x_1) = T(x_2) &\implies x_1 = x_2 \quad \forall x_1, x_2 \in V \end{aligned}$$

o que caracteriza T como biunívoca. Portanto, uma afirmação implica nas outras, o que significa que são equivalentes. ■

Teorema 1.27. *Seja a transformação linear $T : V \mapsto W$, $T \in \mathcal{L}(V, W)$ e $\dim V = n$. Então são afirmações equivalentes:*

1. *T é biunívoca em V.*
2. *Se o conjunto $\{e_i, 1 \leq i \leq k\}$ é independente, então o conjunto $\{T(e_i), 1 \leq i \leq k\}$ é independente.*
3. $\dim T(V) = n$
4. *Se o conjunto $\{e_i, 1 \leq i \leq n\}$ é uma base de V, então o conjunto $\{T(e_i), 1 \leq i \leq n\}$ é uma base de $T(V)$. Em outras palavras, T leva uma base de V para uma base de $T(V)$*

Demonstração. Provando (*1*) $\implies$ (*2*). Seja c_i os escalares (desconhecidos) tais que

$$\begin{aligned} \sum_{i=1}^{k} c_i T(e_i) &= 0 \\ T\left(\sum_{i=1}^{k} c_i e_i\right) &= 0 \end{aligned}$$

Pelo teorema (1.26), T biunívoca implica $N(T) = \{0\}$, ou seja

$$\sum_{i=1}^{k} c_i e_i = 0$$

Por hipótese, o conjunto $\{e_i, 1 \leq i \leq k\}$ é independente, o que significa que $c_i = 0,\ \forall i,\ 1 \leq i \leq k$, e portanto o conjunto $\{T(e_i), 1 \leq i \leq k\}$ também é independente.

Provando *(2)* $\implies$ *(3)*. Agora suponha que se o conjunto $\{e_i, 1 \leq i \leq n\}$ é independente, $\{T(e_i), 1 \leq i \leq n\}$ também é independente, com $k = n$. Suponha também que $T(x) = 0$ para um certo $x \in V$. Pelo teorema (1.7) e sabendo que $\dim V = n$, então o conjunto $\{e_i, 1 \leq i \leq n\}$ é uma base de V e logo x pode ser escrito de tal maneira:

$$x = \sum_{i=1}^{n} d_i e_i$$

logo

$$T(x) = 0 \implies T\left(\sum_{i=1}^{n} d_i e_i\right) = 0 \implies \sum_{i=1}^{n} d_i T(e_i) = 0$$

Se o conjunto $\{e_i, 1 \leq i \leq n\}$ é independente, por hipótese $\{T(e_i), 1 \leq i \leq n\}$ também é independente, no que implica $d_i = 0,\ \forall i,\ 1 \leq i \leq n$. Mas isso implica que $x = 0$ se $T(x) = 0$. Ou seja, $N(T) = \{0\}$ e $\dim N(T) = 0$. Sabendo que V tem dimensão finita, então pela equação (46) do teorema (1.19) temos que

$$\dim T(V) = \dim V = n$$

Provando *(3)* $\implies$ *(4)*. Agora suponha que $\dim T(V) = n$ para um espaço vetorial V de dimensão n e que $A = \{e_i, 1 \leq i \leq n\}$ é uma base de V. Então qualquer base de $T(V)$ tem n elementos. Seja o conjunto $B = \{T(e_i), 1 \leq i \leq n\} \subseteq T(V)$. Sejam os n escalares d_i tais que

$$\sum_{i=1}^{n} d_i T(e_i) = 0 \implies T\left(\sum_{i=1}^{k} d_i e_i\right) = 0$$

Como V tem dimensão finita e $\dim T(V) = n$, então $\dim N(T) = 0$, ou seja, $N(T) = \{0\}$, no que implica

$$\sum_{i=1}^{k} d_i e_i = 0$$

Mas o conjunto A é uma base, e portanto é independente, no que $d_i = 0,\ 1 \leq i \leq n$, ou seja, B é independente. B também tem n elementos, então B é uma base de $T(V)$.

Provando (*4*) $\implies$ (*1*). Suponha que um conjunto $A = \{e_i, 1 \leq i \leq n\}$ ser uma base de V implica que o conjunto $B = \{T(e_i), 1 \leq i \leq n\}$ também é uma base de $T(V)$. Suponha que para algum x_1 e x_2 de V a relação a seguir é válida:

$$T(x_1) = T(x_2) \tag{48}$$

Podemos escrever x_1 e x_2 da seguinte forma:

$$x_1 = \sum_{i=1}^{n} a_i e_i \qquad x_2 = \sum_{i=1}^{n} b_i e_i$$

Aplicando na equação (48), obtemos

$$\sum_{i=1}^{n} a_i T(e_i) = \sum_{i=1}^{n} b_i T(e_i)$$

$$\sum_{i=1}^{n} (a_i - b_i) T(e_i) = 0$$

Mas se A é uma base de V, então por hipótese B também é uma base, em particular é independente. Isso significa que $a_i = b_i$, $\forall i$, $1 \leq i \leq n$. Pelo corolário do teorema (1.7), isso implica que $x_1 = x_2$. Portanto, $T(x_1) = T(x_2) \implies x_1 = x_2$, o que significa que T é biunívoca em V.

■

1.5.5 Transformações lineares no formato matricial

Seja uma C matriz $C = AB$, onde A e B são matrizes de ordem n. Em todo o texto, a multiplicação de matrizes será o produto escalar das linhas A com as colunas de B, ou seja

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

Teorema 1.28. *Dado duas matrizes A e B é válido*

$$(AB)_i = AB_i$$

onde B_i é a i-ésima coluna de B e $(AB)_i$ é a i-ésima coluna de AB.

Demonstração.

$$AB_i = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \cdot \begin{bmatrix} b_{1i} \\ b_{2i} \\ \vdots \\ b_{ni} \end{bmatrix} = \begin{bmatrix} c_{1i} \\ c_{2i} \\ \vdots \\ c_{ni} \end{bmatrix} = (AB)_i$$

■

Seja $T : M_n \mapsto M_n$ a transformação linear

$$T(A) = BA$$

onde M_n é o espaço vetorial das matrizes de ordem n e $A, B \in M_n$. Usando a convenção da multiplicação de matrizes:

$$(BA)_{ij} = \langle \overline{B_i^T}, A_j \rangle$$

onde B_i^T é a i-ésima coluna da transposta de B (ou, em outras palabras, a linha i de B) e o produto interno é definido da seguinte forma:

$$\langle \overline{B_i^T}, A_j \rangle = \sum_{k=1}^{n} b_{ik} a_{kj}$$

Então podemos escrever a i-ésima coluna de BA como

$$(BA)_i = \sum_{k=1}^{n} \langle \overline{B_k^T}, A_i \rangle \cdot I_k$$

Com isso, defina $S : V_n \mapsto V_n$ a transformação linear

$$S(A_i) = BA_i \tag{49}$$

onde V_n é o espaço das matrizes colunas de n elementos, $A_i \in V_n$ e B uma matriz quadrada de ordem n. Pelo teorema (1.28), $S(A_i) = (BA)_i$ e portanto

$$S(A_i) = \sum_{k=1}^{n} \langle \overline{B_k^T}, A_i \rangle \cdot I_k$$

$$T(A) = (S(A_1), S(A_2), \ldots, S(A_n)) \tag{50}$$

Em particular:

$$T(I) = (S(I_1), S(I_2), \ldots, S(I_n)) = BI = B \tag{51}$$

Isso significa que se conhecermos todos os elementos $S(I_i)$ com $1 \leq i \leq n$, então obteremos a matriz B e, consequentemente, a transformação linear T fica totalmente conhecida.

Mais do que isso, qualquer transformação linear pode ter uma representação matricial dada pela equação (49).

Pelo teorema (1.26), se a transformação linear $T : V \mapsto V$ é biunívoca em V, então T possui uma inversa T^{-1}, e esta também tem uma representação matricial. Definamos $T(A) = BA$ e $T^{-1}(C) = DC$, onde A e C denotam as colunas contendo em ordem as coordenadas dos elementos a e c respectivamente (equivalentes matriciais). Então

$$\begin{aligned} T(A) &= BA \\ T^{-1}(T(A)) &= T^{-1}(BA) \\ A &= DBA \end{aligned}$$

Esse resultado só é possível se $DB = I$. Analogamente, podemos realizar o mesmo cálculo, porém fazendo uso de que a imagem de T é o seu próprio domínio:

$$\begin{aligned} T^{-1}(A) &= DA \\ T(T^{-1}(A)) &= T(DA) \\ A &= BDA \end{aligned}$$

Isso só é possível se $BD = I$. E então está provado que

$$B^{-1} = D \quad \text{ou equivalentemente} \quad D^{-1} = B$$

E então podemos enunciar o próximo teorema.

Teorema 1.29. *Dada uma transformação linear T biunívoca cuja imagem é o seu próprio domínio, a matriz que a representa é a inversa da matriz que representa sua transformação inversa T^{-1}.*

Exemplo 1.13 (Mudança de base)**.** *Seja $I_\xi^\beta : V \mapsto V$ uma transformação linear que leva as coordenadas de um elemento $x \in V$ relativas à base β para as coordenadas relativas à ξ. Como I_ξ^β leva uma base de V para outra base de V, pelo teorema (1.27) I_ξ^β é biunívoca e portanto tem uma inversa. Pela representação matricial:*

$$I_\xi^\beta(x) = [I]_\xi^\beta x$$

Agora seja $I_\beta^\xi : V \mapsto V$ que leva as coordenadas relativas à ξ para β:

$$I_\beta^\xi(x) = [I]_\beta^\xi x$$

Por definição, isso significa que $x = [I]_\beta^\xi [I]_\xi^\beta x$. Analogamente, $x = [I]_\xi^\beta [I]_\beta^\xi x$ e portanto $([I]_\beta^\xi)^{-1} = [I]_\xi^\beta$.

1.6 Determinantes

Uma matriz $M_{m \times n}$ pode ser biunivocamente reinterpretada como uma n-upla de matrizes colunas A_i de m elementos cada (ou semelhantemente n-upla de m-uplas). Ou seja,

$$M = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix} = (A_1, A_2, A_3, \dots, A_n)$$

onde

$$A_i = \begin{bmatrix} a_{1i} \\ a_{2i} \\ a_{3i} \\ \vdots \\ a_{mi} \end{bmatrix}, \ \forall i \quad 1 \leq i \leq n$$

Nesta seção, todos os elementos A_i estão ordenados na n-upla de modo que seus subíndices i estão na ordem crescente da esquerda para a direita, exceto quando alguma coluna for permutada.

Definição 1.17. *Seja uma função* det : $M_n \mapsto \mathbb{C}$, *onde* M_n *é o conjunto das matrizes de ordem* n, $M_{n \times n}$. *Denominamos* det *de determinante se satisfazer os seguintes axiomas:*

Axioma 15 (Homogeneidade)**.**

$$\det(\dots, tA_k, \dots) = t \det(\dots, A_k, \dots) \quad \forall t \in \mathbb{C} \quad \forall k \ 1 \leq k \leq n$$

Axioma 16 (Aditividade)**.** *Para toda matriz coluna* C *de* n *elementos:*

$$\det(\dots, A_k + C, \dots) = \det(\dots, A_k, \dots) + \det(\dots, C, \dots) \quad \forall k \ 1 \leq k \leq n$$

Axioma 17 (Dependência)**.**

$$\det(A_1, \dots, A_n) = 0 \quad \textit{se} \quad \exists k \mid A_k = A_{k+1}$$

Axioma 18 (Normalização)**.** *O determinante da matriz identidade* I *é* 1, *ou seja*

$$\det I = 1, \ \textit{onde} \ I = (I_1, \dots, I_n), \ I_i = \begin{bmatrix} \delta_{1i} \\ \delta_{2i} \\ \delta_{3i} \\ \vdots \\ \delta_{ni} \end{bmatrix} \quad \forall i \ 1 \leq i \leq n$$

$$\delta_{ij} = \begin{cases} 0, & \textit{se} \ i \neq j \\ 1, & \textit{se} \ i = j \end{cases}$$

Teorema 1.30. *Uma função* det *satisfazendo os axiomas de* (15) *a* (17) *também satisfaz:*

1. $\det(A_1, \ldots, A_n) = 0$ *se* $A_k = 0$ *para algum* k

2. $\det(\ldots, A_k, A_{k+1}, \ldots) = -\det(\ldots, A_{k+1}, A_k, \ldots) \quad \forall k \mid 1 \leq k < n$

3. $\det(\ldots, A_k, \ldots, A_r \ldots) = -\det(\ldots, A_r, \ldots, A_k \ldots) \quad \forall k \neq r$

4. $\det(A_1, \ldots, A_n) = 0 \quad$ *se* $\quad \exists(i,j),\ i \neq j \mid A_i = A_j$

5. $\det(\ldots, A_k, \ldots) = 0$ *se* $A_k = \sum_{i \neq k} c_i A_i$

Demonstração. (Primeiro item) Seja $M = (\ldots, A_k, \ldots)$ com $A_k = 0$, a matriz coluna nula. Seja também A'_k uma matriz coluna qualquer. Então, pelo axioma (16) da aditividade

$$\det(\ldots, A'_k + 0, \ldots) = \det(\ldots, A'_k, \ldots) + \underbrace{\det(\ldots, 0, \ldots)}_{\det M}$$

$$\det M = \det(\ldots, A'_k, \ldots) - \det(\ldots, A'_k, \ldots) = 0$$

(Segundo item) Sejam $M = (\ldots, A_k, A_{k+1}, \ldots)$ e $M' = (\ldots, A_{k+1}, A_k \ldots)$. Então

$$\begin{aligned}\det M + \det M' &= \det(\ldots, A_k, A_{k+1}, \ldots) + \det(\ldots, A_{k+1}, A_k \ldots) \\ &= \det(\ldots, A_k + A_{k+1}, A_k + A_{k+1}, \ldots) \text{ (axioma 16)} \\ &= 0 \text{ (axioma 17)}\end{aligned}$$

Logo

$$\det M = -\det M' \longrightarrow \det(\ldots, A_k, A_{k+1}, \ldots) = -\det(\ldots, A_{k+1}, A_k, \ldots)$$

(Terceiro item) Seja a matriz $(\ldots, A_k, \ldots, A_r \ldots)$. Permuta-se cada coluna vizinha à esquerda de A_r com ele próprio e percorrendo até trocar de posição com A_k. Cada troca muda o sinal do determinante (item 2 do teorema 1.30) e, portanto, o sinal do determinante no final das trocas será igual a $(-1)^{r-k}$. Para o elemento A_r chegar à antiga posição de A_k, é necessário trocar $r - k - 1$ vezes, de modo que o determinante com as duas colunas trocadas é de $(-1)^{r-k}(-1)^{r-k-1} = (-1)^{2r-2k-1} = (-1)^{2(r-k)-1}$. Mas a potência $2(r-k) - 1$ sempre será um número ímpar, independentemente de r e k. Portanto, $\det(\ldots, A_r, \ldots, A_k \ldots) = -\det(\ldots, A_k, \ldots, A_r \ldots)$.

(Quarto item) Seja M uma matriz que possui pelo menos duas colunas iguais. Trocando essas colunas de lugar, a matriz não muda, porém o determinante troca de sinal (item 3 do teorema 1.30), de modo que

$$\det M = -\det M \implies \det M = 0$$

(Quinto item) Seja a matriz $(\ldots, A_k, \ldots)$, onde $A_k = \sum_{i \neq k} c_i A_i$. Então

$$\begin{aligned}
\det(\ldots, A_k, \ldots) &= (\ldots, \sum_{i \neq k} c_i A_i, \ldots) \\
&= \sum_{i \neq k} \det(\ldots, c_i A_i, \ldots) \\
&= \sum_{i \neq k} c_i \det(\ldots, A_i, \ldots) \\
&= \sum_{i \neq k} c_i \cdot 0 \quad \text{(item 4 do teorema 1.30)} \\
&= 0
\end{aligned}$$

uma vez que para cada matriz $(\ldots, A_i, \ldots)$ há duas colunas A_i iguais.

■

1.6.1 Cálculo de determinantes

Exemplo 1.14. *Seja $M = (A_1, A_2)$, com $A_1 = (a_{11}, a_{21})$ e $A_2 = (a_{12}, a_{22})$. Podemos reescrever essas matrizes colunas da seguinte forma:*

$$A_1 = a_{11} \underbrace{(1,0)}_{\mathbf{i}} + a_{21} \underbrace{(0,1)}_{\mathbf{j}} \qquad A_2 = a_{12}(1,0) + a_{22}(0,1)$$

Então

$$\begin{aligned}
\det M &= \det(A_1, A_2) = \det(a_{11}\mathbf{i} + a_{21}\mathbf{j}; a_{12}\mathbf{i} + a_{22}\mathbf{j}) \\
&= \det(a_{11}\mathbf{i} + a_{21}\mathbf{j}; a_{12}\mathbf{i}) + \det(a_{11}\mathbf{i} + a_{21}\mathbf{j}; a_{22}\mathbf{j}) \\
&= \det(a_{11}\mathbf{i}; a_{12}\mathbf{i}) + \det(a_{21}\mathbf{j}; a_{12}\mathbf{i}) + \det(a_{11}\mathbf{i}; a_{22}\mathbf{j}) + \det(a_{21}\mathbf{j}; a_{22}\mathbf{j}) \\
&= a_{11}a_{12} \det(\mathbf{i}; \mathbf{i}) + a_{21}a_{12} \det(\mathbf{j}; \mathbf{i}) + a_{11}a_{22} \det(\mathbf{i}; \mathbf{j}) + a_{21}a_{22} \det(\mathbf{j}; \mathbf{j})
\end{aligned}$$

Note ainda que

$$\begin{aligned}
&\det(\mathbf{i}; \mathbf{j}) = 1 \ \textit{(axioma 18)} \\
&\det(\mathbf{j}; \mathbf{i}) = -1 \ \textit{(item 3 do teorema 1.30)} \\
&\det(\mathbf{i}; \mathbf{i}) = \det(\mathbf{j}; \mathbf{j}) = 0 \ \textit{(axioma 17)}
\end{aligned}$$

Logo

$$\det M = a_{11}a_{22} - a_{21}a_{12} \tag{52}$$

Exemplo 1.15. *Seja M a matriz diagonal*

$$M = \begin{bmatrix} a_{11} & 0 & 0 & \dots & 0 \\ 0 & a_{22} & 0 & \dots & 0 \\ 0 & 0 & a_{33} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & a_{nn} \end{bmatrix}$$

$$\begin{aligned} \det M &= \det(a_{11}I_1, a_{22}I_2, a_{33}I_3, \dots, a_{nn}I_n) \\ &= a_{11} \cdot a_{22} \cdot a_{33} \cdot \ldots \cdot a_{nn} \det(I_1, I_2, I_3, \dots, I_n) \text{ (axioma 15)} \\ &= a_{11} \cdot a_{22} \cdot a_{33} \cdot \ldots \cdot a_{nn} \text{ (axioma 18)} \end{aligned}$$

Exemplo 1.16. *Seja M a matriz triangular superior*

$$M = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & a_{nn} \end{bmatrix}$$

$$A_j = \sum_{i=1}^{j} a_{ij} I_i$$

Vendo as duas primeiras colunas de M, podemos encontrar uma propriedade das matrizes superiores:

$$\begin{aligned} \det M &= \det(a_{11}I_1, \quad a_{12}I_1 + a_{22}I_2, \dots) \\ &= \det(a_{11}I_1, a_{12}I_1, \dots) + \det(a_{11}I_1, a_{22}I_2, \dots) \\ &= \det(a_{11}I_1, a_{22}I_2, \dots) \end{aligned}$$

Suponha que para um certo $j \geq 1$ seja válido

$$\det\left(\dots, \sum_{i=1}^{j} a_{ij} I_i, \quad \sum_{k=1}^{j+1} a_{k(j+1)} I_k, \dots\right) = \det(\dots, a_{jj}I_j, \quad a_{(j+1)(j+1)}I_{j+1}, \dots)$$

onde as colunas A_i anteriores a A_j no determinante da segunda expressão são iguais a $a_{ii}A_i$. Então

$$\det M = \det\left(\ldots, \sum_{i=1}^{j} a_{ij} I_i, \sum_{i=1}^{j+1} a_{i(j+1)} I_i, \quad \sum_{k=1}^{j+2} a_{k(j+2)} I_k, \ldots\right)$$
$$= \det\left(\ldots, a_{jj} I_j, \quad a_{(j+1)(j+1)} I_{j+1}, \sum_{k=1}^{j+2} a_{k(j+2)} I_k, \ldots\right)$$
$$= \sum_{k=1}^{j+2} a_{k(j+2)} \cdot \det\big(\ldots, a_{jj} I_j, \quad a_{(j+1)(j+1)} I_{j+1}, I_k, \ldots\big)$$

Mas para todos os termos em que $k < j+2$, há alguma coluna de subíndice menor que $j+2$ que é múltipla de I_k, sobrando apenas o termo em que $k = j+2$. Portanto

$$\det M = \det\big(\ldots, a_{jj} I_j, \quad a_{(j+1)(j+1)} I_{j+1}, \quad a_{(j+2)(j+2)} I_{j+2}, \ldots\big)$$

Pelo princípio da indução, foi provado que a matriz M com as colunas

$$A_j = \sum_{i=1}^{j} a_{ij} I_i$$

possui o mesmo determinante da matriz diagonal com as colunas

$$B_i = a_{ii} I_i$$

Logo $\det M = a_{11} \cdot a_{22} \cdot a_{33} \cdot \ldots \cdot a_{nn}$.

Isso mostra também que o determinante não se altera caso alguma coluna seja somada por alguma múltipla de outra.

Exemplo 1.17 (Algoritmo de Gauss-Jordan)**.** *Dado uma matriz M qualquer, podemos transformá-la em uma matriz diagonal superior U usando as operações a seguir.*

1. *Permutar colunas;*
2. *Multiplicar uma coluna por um escalar diferente de zero;*
3. *Adicionar uma coluna ao múltiplo de uma outra.*

A operação (1) altera apenas o sinal do determinante, enquanto (2) multiplica o determinante pelo mesmo escalar e (3) não altera o determinante. Com base nisso, podemos calcular o determinante de uma matriz qualquer com a equação

$$\det M = (-1)^p \cdot (c_1 \cdot \ldots \cdot c_k)^{-1} \det U \tag{53}$$

onde p é o número de permutações de (1), c_i são os escalares multiplicados de (2) e U é a matriz diagonal superior resultante do algoritmo de Gauss-Jordan.

O determinante $\det U$ vale $u_{11}u_{22}\cdot\ldots\cdot u_{nn}\cdot\det I$. Portanto, juntando isso à equação (53), para cada função f que obedece apenas os três primeiros axiomas do determinante existe algum escalar c tal que a equação

$$f(A_1,\ldots,A_n)=c\cdot f(I_1,\ldots,I_n) \tag{54}$$

seja satisfeita para a matriz $(A_1,\ldots,A_n)$.

Teorema 1.31 (Unicidade da função determinante)**.** *Seja* det *uma função obedecendo aos axiomas 15, 16, 17 e 18 e f uma função obedecendo apenas 15, 16 e 17. Então para toda e qualquer matriz $M=(A_1,\ldots,A_n)$ é válido*

$$f(A_1,\ldots,A_n)=\det(A_1,\ldots,A_n)\cdot f(I_1,\ldots,I_n) \tag{55}$$

Ou seja, se f obedecer ao axioma 18, então

$$f(A_1,\ldots,A_n)=\det(A_1,\ldots,A_n)$$

Demonstração. Para mostrarmos a igualdade em (55) basta mostrar que uma função g definida como

$$g(A_1,\ldots,A_n)=f(A_1,\ldots,A_n)-\det(A_1,\ldots,A_n)\cdot f(I_1,\ldots,I_n)$$

é nula para qualquer matriz $(A_1,\ldots,A_n)$.

A função g herda os axiomas 15, 16 e 17 de f e det:

$$\begin{aligned}g(\ldots,tA_k,\ldots)&=f(\ldots,tA_k,\ldots)-\det(\ldots,tA_k,\ldots)\cdot f(I_1,\ldots,I_n)\\&=tf(\ldots,tA_k,\ldots)-t\det(\ldots,A_k,\ldots)\cdot f(I_1,\ldots,I_n)\\&=tg(\ldots,A_k,\ldots)\end{aligned}$$

$$\begin{aligned}g(\ldots,A_k+C,\ldots)&=f(\ldots,A_k+C,\ldots)-\det(\ldots,A_k+C,\ldots)\cdot f(I_1,\ldots,I_n)\\&=f(\ldots,A_k,\ldots)-\det(\ldots,A_k,\ldots)\cdot f(I_1,\ldots,I_n)+\\&+f(\ldots,C,\ldots)-\det(\ldots,C,\ldots)\cdot f(I_1,\ldots,I_n)\\&=g(\ldots,A_k,\ldots)+g(\ldots,C,\ldots)\end{aligned}$$

$$\begin{aligned}\det(\ldots,A_k,A_k,\ldots)&=f(\ldots,A_k,A_k,\ldots)-\det(\ldots,A_k,A_k,\ldots)\cdot f(I_1,\ldots,I_n)\\&=0-0\cdot f(I_1,\ldots,I_n)\\&=0\end{aligned}$$

Portanto, g obedece à equação (54):

$$g(A_1,\ldots,A_n)=c\cdot g(I_1,\ldots,I_n)$$

Usando a definição de g:

$$\begin{aligned} g(I_1, \ldots, I_n) &= f(I_1, \ldots, I_n) - \det(I_1, \ldots, I_n) \cdot f(I_1, \ldots, I_n) \\ &= f(I_1, \ldots, I_n) - 1 \cdot f(I_1, \ldots, I_n) \\ &= 0 \end{aligned}$$

e então

$$g(A_1, \ldots, A_n) = c \cdot 0 = 0$$

Portanto

$$f(A_1, \ldots, A_n) = \det(A_1, \ldots, A_n) \cdot f(I_1, \ldots, I_n)$$

para qualquer matriz $(A_1, \ldots, A_n)$.

■

Como resultado disso, segue que a função det é única.

1.6.2 Produto de determinantes

Teorema 1.32 (Produto de determinantes)**.** *Se A e B são duas matrizes de ordem n, então*

$$\det(AB) = (\det A)(\det B)$$

Demonstração. Usando o teorema (1.28):

$$\det((AB)_1, \ldots, (AB)_n) = \det(AB_1, \ldots, AB_n)$$

Seja f_A a função

$$f_A(B_1, \ldots, B_n) = \det(AB_1, \ldots, AB_n)$$

No caso particular em que $B = I$, a matriz identidade de ordem n:

$$f_A(I_1, \ldots, I_n) = \det(AI_1, \ldots, AI_n) = \det A$$

A proposição $\det(AB) = (\det A)(\det B)$ está correta se e somente se a equação seguinte for válida:

$$\det(AB_1, \ldots, AB_n) = \det(B_1, \ldots, B_n) \cdot f_A(I_1, \ldots, I_n) \tag{56}$$

A equação (56) pode ser reescrita como

$$f_A(B_1, \ldots, B_n) = \det(B_1, \ldots, B_n) \cdot f_A(I_1, \ldots, I_n) \tag{57}$$

Os axiomas (15) a (17) são satisfeitos por f:

$$
\begin{aligned}
f_A(\ldots, tB_i, \ldots) &= \det(\ldots, AtB_i, \ldots) \\
&= t\det(\ldots, AB_i, \ldots) \\
&= tf_A(\ldots, B_i, \ldots)
\end{aligned}
$$

$$
\begin{aligned}
f_A(\ldots, B_i + C, \ldots) &= \det(\ldots, A(B_i + C), \ldots) \\
&= \det(\ldots, AB_i + AC, \ldots) \\
&= \det(\ldots, AB_i, \ldots) + \det(\ldots, AC, \ldots) \\
&= f_A(\ldots, B_i, \ldots) + f_A(\ldots, C, \ldots)
\end{aligned}
$$

$$
\begin{aligned}
f_A(\ldots, B_i, B_i, \ldots) &= \det(\ldots, AB_i, AB_i, \ldots) \\
&= 0
\end{aligned}
$$

Pelo teorema (1.31) da unicidade, a função f_A obedece a equação (57) para qualquer matriz $B = (B_1, \ldots, B_n)$. Segue então que

$$\det(AB) = (\det A)(\det B)$$

■

Note que, como os números complexos são comutativos na multiplicação, é rápido verificar que o determinante de AB é igual ao de BA.

Corolário (Comutatividade do determinante no corpo dos complexos)**.** *Para quaisquer matrizes A e B de ordem n é válido*

$$\det(AB) = \det(BA)$$

Teorema 1.33 (Determinante da matriz inversa)**.** *Para qualquer matriz A que* $\det A \neq 0$ *e possui uma inversa A^{-1} é válido*

$$\det\left(A^{-1}\right) = \frac{1}{\det(A)}$$

onde A^{-1} é a matriz tal que $AA^{-1} = A^{-1}A = I$

Demonstração.

$$
\begin{aligned}
\det(I) = \det\left(AA^{-1}\right) &= \det(A)\det\left(A^{-1}\right) = 1 \\
\det\left(A^{-1}\right) &= \frac{1}{\det(A)}
\end{aligned}
$$

■

Posteriormente será provado que $\det A \neq 0$ e existir inversa A^{-1} são afirmações equivalentes.

1.6.3 Independência entre colunas

A propriedade *5* do teorema (1.30) nos diz que se uma matriz M possui pelo menos uma coluna que é combinação linear de outras colunas, então o seu determinante é zero. A contrapositiva dessa proposição nos diz que se $\det M \neq 0$ então não há colunas dependentes e, portanto, suas colunas são independentes. Mas isso ainda não prova se uma matriz M tiver colunas independentes implica em $\det M \neq 0$.

Teorema 1.34 (Determinante e Independência)**.** *As colunas de uma matriz M de ordem n são independentes entre si se e somente se* $\det M \neq 0$

Demonstração. A proposição $\det M \neq 0 \implies$ independência das colunas de M já está provada pelo teorema (1.30).

Se $A = (A_1, \ldots, A_n)$ não possui colunas dependentes entre si, então

$$\sum_{i=1}^{n} d_i A_i = 0 \implies d_i = 0,\ \forall i$$

e os elementos A_i formam uma base para o espaço gerado por eles, V_n.

Seja T uma transformação linear que satisfaça

$$T(A_i) = I_i,\ \forall i$$

Como visto na seção 1.5.5, T tem uma representação matricial:

$$T(A_i) = BA_i$$

Mas $T(A_i) = I_i$, e isso significa que

$$\begin{aligned} BA_i &= I_i \\ BA &= I \\ \det(BA) &= \det(I) \\ \det(B) \cdot \det(A) &= 1 \end{aligned}$$

Por conseguinte, $\det(A) \neq 0$

■

Note que no caso acima, B é a matriz inversa de A.

Teorema 1.35 (Determinante e T^{-1})**.** *Seja $T : V \mapsto V$ uma transformação linear. Então o determinante da matriz que representa T é diferente de zero se e somente se T for biunívoca em V. Equivalentemente, a matriz $[T]$ possui uma inversa $[T]^{-1}$ se e somente se T for biunívoca em V.*

Demonstração. Se T é biunívoca, então T possui uma inversa T^{-1} e, consequentemente, na representação matricial

$$[T][T^{-1}] = [T^{-1}][T] = I$$

Aplicando o determinante nessa equação, obtemos

$$\det\big([T][T^{-1}]\big) = \det(I)$$
$$\det([T]) \cdot \det\big([T^{-1}]\big) = 1$$

então $\det([T]) \neq 0$ e $\det\big([T^{-1}]\big) \neq 0$.

Agora suponha que $\det[T] \neq 0$. Seja X uma matriz coluna representando o elemento x tal que $T(x) = 0$:

$$[T]X = 0 \quad \text{com } X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Usando a convenção adotada da multiplicação de matrizes, podemos reescrever a equação $[T]X = 0$ da seguinte forma:

$$x_1[T]_1 + x_2[T]_2 + \ldots + x_n[T]_n = 0$$

onde $[T]_i$ é a i-ésima coluna de $[T]$ e 0 é a coluna nula.

Por hipótese, $\det[T] \neq 0$ e então, pelo teorema (1.34), as colunas de $[T]$ são linearmente independentes entre si, o que significa que $\forall i,\ x_i = 0$. E então que $T(x) = 0 \implies x = 0$, no que implica $N(T) = \{0\}$. Portanto, pelo Item 3 do teorema (1.26), T é biunívoca e possui uma transformação inversa T^{-1} cuja representação matricial é a inversa de $[T]$.

■

O próximo teorema é um corolário importante no estudo de equações matriciais homogêneas.

Teorema 1.36. *Dada uma equação matricial $Ax = 0$, o sistema possui mais de uma solução se e somente se* $\det A = 0$. *Equivalentemente, $Ax = 0$ só possui uma solução se e somente se* $\det A \neq 0$.

Demonstração. Seja T a transformação linear representada pela matriz A. Suponha que $N(T) = \{0\}$. Pelo teorema (1.26), T é biunívoca e então o teorema (1.35) garante que o determinante de A é diferente de zero.

Agora suponha que a equação matricial tenha mais de uma solução. Então T não é biunívoca e o teorema (1.35) nos garante que $\det A = 0$. ■

O próximo teorema é uma adaptação para sistemas não-homogêneos, ou seja, para quaisquer sistemas do tipo $Ax = b$, com b não necessariamente sendo uma matriz nula.

Teorema 1.37. $\det A \neq 0$ *se e somente se o sistema* $Ax = b$ *possuir apenas uma solução.*

Demonstração. Se x_1 e x_2 são soluções de $Ax = b$, então

$$\begin{aligned} Ax_1 - Ax_2 &= A(x_1 - x_2) \\ b - b &= A(x_1 - x_2) \\ 0 &= A(x_1 - x_2) \end{aligned}$$

Se $\det A \neq 0$, o teorema (1.36) garante que a equação homogênea $A(x_1 - x_2) = 0$ só possui uma solução, que é a trivial. Isso significa que $x_1 = x_2$. Além disso, a solução de $Ax = b$ é dada por $x = A^{-1}b$, uma vez que o teorema (1.35) garante que A possui uma inversa.

E no caso de $Ax = b$ possuir apenas uma solução, então o sistema homogêneo $Ay = 0$ só pode ter uma única solução (a trivial), pois caso contrário

$$A(x + y) = Ax + Ay = b + 0 = b$$

com $x + y \neq x$, o que violaria a hipótese de solução única de $Ax = b$. Então, pelo teorema (1.36), $\det A \neq 0$. ■

1.6.4 Determinantes de blocos de matrizes

Algumas matrizes possuem um formato especial de maneira que podemos reescrever suas entradas como matrizes.

Exemplo 1.18 (Blocos diagonais)**.** *Um exemplo de matriz diagonal por blocos de ordem* 2 *é a matriz seguinte:*

$$C = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 & 0 \\ 0 & 0 & b_{11} & b_{12} & b_{13} \\ 0 & 0 & b_{21} & b_{22} & b_{23} \\ 0 & 0 & b_{31} & b_{32} & b_{33} \end{bmatrix}$$

onde

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}$$

e 0 *são as matrizes de zeros que completam a matriz quadrada* C*.*

Teorema 1.38. *Sejam* A *e* B *duas matrizes quadradas quaisquer. Então*

$$\det \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} = (\det A)(\det B)$$

Demonstração. Suponha que a matriz A tenha ordem n e que B tenha ordem m. Então

$$\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & I^m \end{bmatrix} \cdot \begin{bmatrix} I^n & 0 \\ 0 & B \end{bmatrix}$$

onde I^m e I^n são as matrizes identidades de ordem m e n respectivamente. Logo

$$\det \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} = \det \begin{bmatrix} A & 0 \\ 0 & I^m \end{bmatrix} \cdot \det \begin{bmatrix} I^n & 0 \\ 0 & B \end{bmatrix}$$

Seja f a função

$$f(A) = \det \begin{bmatrix} A & 0 \\ 0 & I^m \end{bmatrix}$$

Verificam-se os axiomas (15), (16) e (17) em f:

$$\begin{aligned} f(\ldots, tA_i, \ldots) &= \det \begin{bmatrix} (\ldots, tA_i, \ldots) & 0 \\ (\ldots, t0_i, \ldots) & (I_1, \ldots, I_m) \end{bmatrix} \\ &= t \det \begin{bmatrix} A & 0 \\ 0 & I^m \end{bmatrix} \\ &= t \cdot f(\ldots, A_i, \ldots) \end{aligned}$$

onde 0_i é a i-ésima matriz coluna de zeros.
Também

$$f(\ldots, A_i + C, \ldots) = \det \begin{bmatrix} (\ldots, A_i + C, \ldots) & 0 \\ (\ldots, 0_i + 0, \ldots) & (I_1, \ldots, I_m) \end{bmatrix}$$

Seja C' a coluna C com outros $m - n$ zeros após o último elemento. Então

$$\begin{aligned} f(\ldots, A_i + C, \ldots) &= \det \begin{bmatrix} (\ldots, A_i + C', \ldots) & 0 \\ (\ldots, 0_i + C', \ldots) & (I_1, \ldots, I_m) \end{bmatrix} \\ &= \det \begin{bmatrix} (\ldots, A_i, \ldots) & 0 \\ (\ldots, 0_i, \ldots) & (I_1, \ldots, I_m) \end{bmatrix} + \\ &+ \det \begin{bmatrix} (\ldots, C', \ldots) & 0 \\ (\ldots, C', \ldots) & (I_1, \ldots, I_m) \end{bmatrix} \\ &= f(\ldots, A_i, \ldots) + f(\ldots, C, \ldots) \end{aligned}$$

E ainda

$$f(\ldots, A_i, A_i, \ldots) = \det \begin{bmatrix} (\ldots, A_i, A_i, \ldots) & 0 \\ (\ldots, 0_i, 0_i, \ldots) & (I_1, \ldots, I_m) \end{bmatrix} = 0$$

Pelo teorema da unicidade, f satisfaz a seguinte equação:

$$f(A_1, \ldots, A_n) = \det(A_1, \ldots, A_n) \cdot f(I_1, \ldots, I_n)$$

Mas

$$f(I_1, \ldots, I_n) = \det \begin{bmatrix} I^n & 0 \\ 0 & I^m \end{bmatrix} = \det I^{m+n} = 1$$

Pela unicidade da função determinante, segue que

$$f(A) = \det \begin{bmatrix} A & 0 \\ 0 & I^m \end{bmatrix} = \det A$$

Com um raciocínio análogo, é possível provar que

$$\det \begin{bmatrix} I^n & 0 \\ 0 & B \end{bmatrix} = \det B$$

Então

$$\det \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} = (\det A)(\det B)$$

■

1.6.5 Formula geral do determinante

Teorema 1.39. *O determinante da transposta A^T de uma matriz é o próprio determinante da matriz A.*

$$\det A^T = \det A \tag{58}$$

Demonstração. Seja T a transformação linear que leva uma matriz à sua transposta:

$$T(A) = A^T$$

A transformação T possui uma representação matricial cujo operador é a multiplicação à esquerda por B, $T(A) = BA$. Usando a fórmula do produto de determinantes do teorema (1.32) com $A = I$ (cuja transposta é a própria identidade), temos

$$\begin{aligned} \det(BI) &= \det I^T \\ \det B \cdot \det I &= \det I \\ \det B &= 1 \end{aligned}$$

Portanto:

$$\begin{aligned}
\det(BA) &= \det A^T \\
\det B \cdot \det A &= \det A^T \\
\det A &= \det A^T
\end{aligned}$$

■

A equação (58) é muito importante, pois nos diz que se as três operações sobre colunas do algoritmo de Gauss-Jordan (exemplo 1.17) forem usadas nas linhas de uma matriz ao invés de suas colunas, as alterações no determinante são idênticas.

Corolário. *Dado uma matriz M qualquer:*

1. *Permutar linhas troca o sinal do determinante;*
2. *Multiplicar uma linha por um escalar diferente de zero multiplica o determinante pelo mesmo escalar;*
3. *Adicionar uma linha ao múltiplo de uma outra não altera o determinante.*

Com esse corolário, podemos elaborar uma fórmula geral do determinante para matrizes de qualquer ordem maior do que 2.

Seja $A = (A_1, \ldots, A_n)$, com $n > 2$. A coluna A_i pode ser escrita como

$$A_i = \sum_{k=1}^{n} a_{ki} I_k$$

e portanto o determinante de A pode ser calculado com a seguinte expressão:

$$\begin{aligned}
\det(\ldots, A_{i-1}, A_i, A_{i+1} \ldots) &= \det\left(\ldots, A_{i-1}, \sum_{k=1}^{n} a_{ki} I_k, , A_{i+1} \ldots\right) \\
&= \sum_{k=1}^{n} a_{ki} \det(\ldots, A_{i-1}, I_k, A_{i+1}, \ldots) \\
&= \sum_{k=1}^{n} L_{ki}
\end{aligned}$$

Pegando o k-ésimo termo do somatório e permutando a coluna I_k com suas vizinhas até chegar a primeira coluna, temos

$$L_{ki} = (-1)^{i-1} a_{ki} \cdot \det(I_k, A_1, \ldots, A_{i-1}, A_{i+1}, \ldots)$$

Não confundir a k-ésima coluna da matriz identidade I_k com a k-ésima coluna da matriz da equação. Houve $i - 1$ permutações.

Permutando a linha k com suas vizinhas até chegar à primeira linha:

$$
\begin{aligned}
L_{ki} &= (-1)^{(i-1)+(k-1)} a_{ki} \cdot \det(I_1, A'_1, A'_2, \ldots, A'_n) \\
&= (-1)^{i+k} a_{ki} \cdot \det(I_1, A'_1, A'_2, \ldots, A'_n)
\end{aligned}
$$

onde A' denota as colunas que tiveram suas linhas trocadas. Usando o axioma (16) como uma operação sobre linhas:

$$
\begin{aligned}
(-1)^{i+k+1} L_{ki} = a_{ki} \cdot 1 \cdot \det\big(I_1, A^0_2, A^0_3, \ldots, A^0_n\big) + \\
+ a_{ki} \cdot a_{12} \cdot \det\big(0, A^1_2, A^0_3, \ldots, A^0_n\big) + \\
+ \ldots + a_{ki} \cdot a_{1n} \cdot \det\big(0, A^0_2, \ldots, A^1_n\big)
\end{aligned}
$$

onde A^0_i e A^1_i denota uma coluna i que possui 0 e 1 respectivamente no seu primeiro elemento. Note que todos os termos são nulos, exceto o primeiro, por conterem a primeira coluna nula. Então:

$$
L_{ki} = (-1)^{i+k} a_{ki} \cdot 1 \cdot \det\big(I_1, A^0_2, A^0_3, \ldots, A^0_n\big)
$$

Essa matriz é diagonal por blocos:

$$
\det\big(I_1, A^0_2, A^0_3, \ldots, A^0_n\big) = \det \begin{bmatrix} 1 & 0 \\ 0 & A^{ki} \end{bmatrix}
$$

onde A^{ki} é a matriz A com coluna i e linha k suprimidos. Pelo teorema (1.38):

$$
\begin{aligned}
L_{ki} &= (-1)^{i+k} a_{ki} \cdot \det[1] \cdot \det A^{ki} \\
&= (-1)^{i+k} a_{ki} \cdot \det A^{ki}
\end{aligned}
$$

Logo:

$$
\det(A_1, \ldots, A_i, \ldots, A_n) = \sum_{k=1}^{n} (-1)^{i+k} a_{ki} \cdot \det A^{ki} \tag{59}
$$

A matriz A^{ki} possui ordem $n - 1$, e isso significa que se repetitivamente reaplicarmos a equação (59) para calcular $\det A^{ik}$, a ordem da matriz inexoravelmente se reduzirá a 2, cujo determinante é dado pela equação (52).

Como o determinante é indiferente entre operações sobre linhas ou sobre colunas, podemos reinterpretar (59) como sendo o determinante da matriz de linhas A_i. Portanto é aceitável trocar os subíndices da equação para tornar possível a escolha de varrer uma linha ao invés de uma coluna:

$$
\det(A_1, \ldots, A_i, \ldots, A_n) = \sum_{k=1}^{n} (-1)^{i+k} a_{ik} \cdot \det A^{ik} \tag{60}
$$

O índice i é arbitrário, então na prática é escolhido uma linha ou coluna que tenha mais zeros (quando possível) para reduzir a quantidade de termos do somatório.

1.6.6 Wronskiana

Definição 1.18. *Sejam n funções $u_j(x)$ n vezes deriváveis e de uma única variável x. A matriz wronskiana $W(x)$, quadrada de ordem n, é definida pela expressão*

$$W(x) = \begin{bmatrix} u_1(x) & u_2(x) & \dots & u_n(x) \\ u_1'(x) & u_2'(x) & \dots & u_n'(x) \\ u_1''(x) & u_2''(x) & \dots & u_n''(x) \\ \vdots & \vdots & \ddots & \vdots \\ u_1^{(n)}(x) & u_2^{(n)}(x) & \dots & u_n^{(n)}(x) \end{bmatrix}$$

Note que a wronskiana pode ser reescrita na forma

$$W^T(x) = \left(U(x), \frac{d}{dx}U(x), \frac{d^2}{dx^2}U(x), \dots, \frac{d^n}{dx^n}U(x) \right)$$

onde

$$U(x) = \begin{bmatrix} u_1(x) \\ u_2(x) \\ \vdots \\ u_n(x) \end{bmatrix}$$

Teorema 1.40. *Seja a matriz $M = [f_{ij}]$ de ordem n, onde f_{ij} é uma função derivável no intervalo $[a, b]$. Se $F = \det M$, então*

$$F' = \sum_{i=1}^{n} \det M_i \tag{61}$$

onde M_i é a matriz obtida ao derivar as funções da linha i de M.

Demonstração. Seja $M = [f_{ij}(x)]$ uma matriz quadrada de ordem 2. Já vimos que

$$\det M = F = f_{11}f_{22} - f_{12}f_{21}$$

Derivando a equação, obtemos

$$\begin{aligned} F' &= f_{11}'f_{22} + f_{11}f_{22}' - f_{12}'f_{21} - f_{12}f_{21}' \\ &= [f_{11}'f_{22} - f_{12}'f_{21}] + [f_{11}f_{22}' - f_{12}f_{21}'] \\ &= \det M_1 + \det M_2 \end{aligned}$$

Portanto o teorema é válido para $n = 2$. Agora suponha que $M = [f_{ij}]$ uma matriz quadrada de ordem n e que o teorema seja válido para matrizes do tipo de ordem $n - 1$.

Usando $i = n$ em (60), a função $F(x)$ pode ser expressa como

$$F = \sum_{k=1}^{n} (-1)^{n+k} f_{nk} \cdot \det M^{nk}$$

Derivando a equação, obtemos

$$F' = \sum_{k=1}^{n} (-1)^{n+k} f'_{nk} \cdot \det M^{nk} + \sum_{k=1}^{n} (-1)^{n+k} f_{nk} (\det M^{nk})'$$

Sabendo que M^{nk} tem ordem $n-1$, então por hipótese:

$$(\det M^{nk})' = \sum_{i=1}^{n-1} \det M_i^{nk}$$

e consequentemente

$$F'(x) = \sum_{k=1}^{n} (-1)^{n+k} f'_{nk} \cdot \det M^{nk} + \sum_{i=1}^{n-1} \sum_{k=1}^{n} (-1)^{n+k} f_{nk} \det M_i^{nk}$$

Note que

$$\sum_{k=1}^{n} (-1)^{n+k} f'_{nk} \cdot \det M^{nk} = \det M_n$$

e também

$$\sum_{k=1}^{n} (-1)^{1+k} f_{1k} \det M_i^{ik} = \det M_i$$

Portanto

$$\begin{aligned} F'(x) &= \det M_n + \sum_{i=1}^{n-1} \det M_i \\ &= \sum_{i=1}^{n} \det M_i \end{aligned}$$

Por indução, a equação (61) é válida para todo o $n \geq 2$.

■

Teorema 1.41 (Derivada do determinante da Wronskiana)**.** *Seja f a função $f(x) = \det W(x)$. Então a derivada $f'(x)$ é o determinante da matriz obtida ao derivar a última linha de $W(x)$.*

Demonstração. Pelo teorema (1.40)

$$\begin{aligned} f'(x) &= \sum_{i=1}^{n} \det W_i(x) \\ &= \det W_1(x) + \det W_2(x) + \ldots + \det W_n(x) \end{aligned}$$

onde W_i é a matriz obtida derivando a linha i de $W(x)$.

Mas todos os termos $W_i(x)$, com $1 \leq i < n$, possuem duas linhas idênticas vizinhas, o que significa que $\det W_i(x) = 0$, exceto para $\det W_n(x)$. Portanto

$$f'(x) = \frac{d}{dx}(\det W(x)) = W_n(x)$$

■

Teorema 1.42 (Wronskiana e equações diferenciais lineares homogêneas)**.** *Seja a equação diferencial*

$$f''(x) + a_1(x)f'(x) + a_0(x)f(x) = 0 \quad \forall x \in I \tag{62}$$

onde a_1 é uma função contínua em I. Então a wronskiana W de duas soluções quaisquer é identicamente nula em I ou $W \neq 0$ para todo x em I.

Demonstração. Como f_1 e f_2 são soluções da equação (62), então

$$\begin{aligned} f_1'' + a_1 f_1' + a_0 f_1 = 0 \\ f_2'' + a_1 f_2' + a_0 f_2 = 0 \end{aligned}$$

Multiplicando a primeira equação por f_2 e a segunda por f_1, obtemos

$$\begin{aligned} f_2 f_1'' + a_1 f_2 f_1' + a_0 f_2 f_1 = 0 \\ f_1 f_2'' + a_1 f_1 f_2' + a_0 f_1 f_2 = 0 \end{aligned}$$

Subtraindo uma equação da outra, temos

$$(f_2 f_1'' - f_1 f_2'') + a_1(f_2 f_1' - f_1 f_2') = 0 \tag{63}$$

Agora note que

$$f_2 f_1'' - f_1 f_2'' = \det \begin{bmatrix} f_2 & f_1 \\ f_2'' & f_1'' \end{bmatrix} \quad \text{e} \quad f_2 f_1' - f_1 f_2' = \det \begin{bmatrix} f_2 & f_1 \\ f_2' & f_1' \end{bmatrix}$$

De acordo com o teorema (1.41), O primeiro determinante é $W'(f_2, f_1)$. O segundo determinante é $W(f_2, f_1)$ e portanto a equação (63) fica

$$W'(x) + a_1(x)W(x) = 0$$

$$W'(x) = -a_1(x)W(x)$$

A solução geral para essa equação é

$$W(x) = c \cdot \exp\left(-\int_a^x a_1(x')\,\mathrm{d}x'\right)$$

onde c é uma constante arbitrária. Visto que a exponencial nunca se anula (a integral não pode tender ao infinito já que $a_1(x)$ é contínua por hipótese), caso $W(x)$ se anule em algum ponto $x_0 \in I$, então $W(x)$ é identicamente nula em I. ■

Teorema 1.43 (Wronskiana e a independência linear de soluções)**.** *Se f_1 e f_2 são soluções da equação diferencial* (62)*, então uma é linearmente independente da outra se e somente se $W(f_1, f_2)(x)$ não for a função identicamente nula em I.*

Demonstração. Se f_1 e f_2 são linearmente dependentes tal que $f_2 = \alpha f_1$, então

$$W(f_1, f_2)(x) = f_1 \alpha f_1' - \alpha f_1 f_1' = 0$$

A contrapositiva é que se $W(f_1, f_2)(x) \neq 0$ em I, então f_1 e f_2 são duas soluções linearmente independentes.

Por outro lado, se $W(x) = 0$ em algum ponto de I, então pelo teorema (1.42) $W(x)$ é a função identicamente nula. Como $W(x)$ é o determinante

$$\det\begin{bmatrix} f_1 & f_2 \\ f_1' & f_2' \end{bmatrix}$$

, o teorema (1.30) garante que os vetores coluna são linearmente dependentes entre si, ou seja, existe algum α tal que $\alpha f_1 = f_2$. A contrapositiva é que se f_1 e f_2 são linearmente independentes, então $W(f_1, f_2)(x) \neq 0$ em I. ■

1.7 Autovalores e Autovetores

Teorema 1.44. *Seja $T : V \mapsto V$ uma transformação linear, com $\dim V = n$. Se existir alguma base $\xi = (u_1, \ldots, u_n)$ de V tal que a representação matricial de T_ξ^ξ seja uma matriz diagonal, então T na base canônica*

$$T(u_i) = \lambda_i u_i \quad \forall i,\ 1 \leq i \leq n \tag{64}$$

Reciprocamente, se existe uma base $\xi = (u_1, \ldots, u_n)$ tal que a equação (64) *seja válida para todos os termos dessa base, então a representação matricial de T_ξ^ξ é uma matriz diagonal.*

Demonstração. A transformação pode ser representada na forma matricial:

$$[T]_\xi^\xi = [I]_\xi^\beta [T]_\beta^\beta [I]_\beta^\xi \tag{65}$$

onde $[T]_\beta^\beta$ é a transformação T na base canônica β, $[I]_\xi^\beta$ representa a matriz mudança de base de β para ξ e analogamente para $[I]_\beta^\xi$. Como já vimos

anteriormente, as matrizes mudança de base possuem uma inversa. Portanto podemos aplicá-las na equação (65):

$$([I]_\xi^\beta)^{-1}[T]_\xi^\xi([I]_\beta^\xi)^{-1} = [T]_\beta^\beta$$

Pelo teorema (1.29), a matriz $([I]_\xi^\beta)^{-1}$ denota a matriz mudança de base inversa, ou seja, de ξ para β. Raciocínio análogo para $([I]_\beta^\xi)^{-1}$. Logo:

$$([I]_\xi^\beta)^{-1} = [I]_\beta^\xi \quad ([I]_\beta^\xi)^{-1} = [I]_\xi^\beta$$

$$[I]_\beta^\xi[T]_\xi^\xi[I]_\xi^\beta = [T]_\beta^\beta$$

Aplicando a transformação em um elemento da base ξ qualquer, u_i, cuja representação matricial é denotado pela matriz coluna U_i, obtemos

$$[I]_\beta^\xi[T]_\xi^\xi[I]_\xi^\beta U_i = [T]_\beta^\beta U_i$$

Mas U_i na base ξ é I_i:

$$[I]_\beta^\xi[T]_\xi^\xi I_i = [T]_\beta^\beta U_i$$

Por hipótese, $[T]_\xi^\xi$ é uma matriz diagonal. Então o produto $[T]_\xi^\xi I_i$ resultará na matriz coluna $\lambda_i \cdot I_i$, onde λ_i é o escalar da linha i e coluna i de $[T]_\xi^\xi$. Portanto

$$\lambda_i[I]_\beta^\xi \cdot I_i = [T]_\beta^\beta U_i$$

Como a matriz $[I]_\xi^\beta$ leva U_i para I_i, então a sua inversa $[I]_\beta^\xi$ leva I_i para U_i. E finalmente

$$\lambda_i U_i = [T]_\beta^\beta U_i$$

ou

$$T(u_i) = \lambda_i u_i$$

Agora, para provar a recíproca, suponha que existe uma base $\xi = (u_1, \ldots, u_n)$ tal que a equação (64) seja válida para todos os termos dessa base. A equação (51) nos diz que

$$S(I) = (T(I_1), T(I_2), \ldots, T(I_n)) = [T]$$

Mas a base não é especificada e consequentemente é arbitrária. Por conveniência, vamos escolher a base ξ:

$$[T]_\xi^\xi = (T_\xi^\xi(I_1), T_\xi^\xi(I_2), \ldots, T_\xi^\xi(I_n))$$

Portanto, para mostrar que $[T]_\xi^\xi$ é diagonal, basta provar que

$$[T]_\xi^\xi I_i = \lambda_i I_i$$

Aplicando a matriz $[T]_\xi^\xi$ na coluna I_i da matriz identidade e novamente usando a equação (65):

$$\begin{aligned}[T]_\xi^\xi I_i &= [I]_\xi^\beta [T]_\beta^\beta [I]_\beta^\xi I_i \\ &= [I]_\xi^\beta [T]_\beta^\beta U_i\end{aligned}$$

Note que $[T]_\beta^\beta U_i$ é a forma matricial de $T(u_i) = \lambda_i u_i$, portanto $[T]_\beta^\beta U_i = \lambda_i U_i$ e

$$\begin{aligned}[T]_\xi^\xi I_i &= [I]_\xi^\beta \lambda_i U_i \\ &= \lambda_i [I]_\xi^\beta U_i \\ &= \lambda_i I_i\end{aligned}$$

o que completa a demonstração.

■

Os escalares λ_i e os elementos u_i que obedecem (64) são especiais e serão analisados com mais profundidade nesta seção.

Definição 1.19. *Seja $T : V \mapsto W$ uma transformação linear com $V \subseteq W$. Denominamos λ de autovalor se existe algum $u \in V,\ u \neq 0$ tal que*

$$T(u) = \lambda u \tag{66}$$

O elemento u correlacionado ao autovalor λ é denominado de autovetor.

Para cada autovetor u de T está associado a apenas um único autovalor λ, pois se $T(u) = \lambda u$ e $T(u) = \gamma u$, então

$$0 = T(u) - T(u) = \lambda u - \gamma u = (\lambda - \gamma)u$$

Como u é um autovetor, então $u \neq 0$ e $\lambda - \gamma = 0 \implies \gamma = \lambda$.

Mas a recíproca não é verdadeira. Existem casos onde para um certo autovalor λ há vários autovetores linearmente independente associados.

Note que o múltiplo de um autovetor, devido à linearidade da transformação, também é um autovetor e ambos estão correlacionados com o mesmo autovalor:

$$T(u) = \lambda u \implies T(au) = aT(u) = a\lambda u = \lambda(au)$$

Exemplo 1.19 (Multiplicação por um escalar)**.** *Seja* $T : V \mapsto W$, $T(x) = cx$, $\forall x \in V$ *e* $\dim V = n$. *Qualquer elemento não-nulo de* V *é um autovetor com autovalor* c. *Mais precisamente há* n *autovetores linearmente independentes e portanto podemos construir uma base de autovetores de* V.

Exemplo 1.20 (Subespaço dos autovetores)**.** *Seja* $E_T(\lambda)$ *o conjunto de todos os elementos* x *(incluindo 0) tais que* $T(x) = \lambda x$. $E_T(\lambda)$ *obedece aos axiomas do subespaço vetorial:*

$$x, y \in E_T(\lambda),\ T(x+y) = T(x) + T(y) = \lambda x + \lambda y = \lambda(x+y)$$
$$x \in E_T,\ T(cx) = cT(x) = c\lambda x = \lambda(cx)$$

onde c *é um escalar. Se* $E_T(\lambda)$ *possui dimensão finita, e* T *possui* λ *como um de seus autovalores, então* $\dim E_T(\lambda) \geq 1$, *uma vez que há pelo menos um autovetor (ou seja, não-nulo).*

Exemplo 1.21 (Nulidade maior que zero e o Autovalor nulo)**.** *Seja* $T : V \mapsto W$ *uma transformação linear cuja nulidade é diferente de zero, ou seja, existe pelo menos um elemento não-nulo* x *de* V *tal que* $T(x) = 0$, *ou melhor,* $T(x) = 0x$. *Transformações lineares como essa possuem o escalar zero como autovalor e o subespaço* $E_T(\lambda = 0)$ *se torna o núcleo de* T. *Em geral, se* $S(x) = T(x) - \lambda x$, *o subespaço* $E_T(\lambda)$ *se torna o núcleo de* S.

Exemplo 1.22 (Reflexão no plano XY)**.** *Seja* $T : \mathbb{R}^3 \mapsto \mathbb{R}^3$ *a transformação que leva o segmento de reta orientado* (x, y, z) *ao* $(x, y, -z)$. *Os elementos* $(x, y, 0)$ *são autovetores de autovalor* 1, *enquanto que os elementos* $(0, 0, z)$ *são autovetores de autovalor* -1. *Os elementos* $(1, 0, 0)$ *e* $(0, 1, 0)$ *da base canônica do* $\mathbb{R}^3$ *formam uma base de* $E_T(1)$ *enquanto que* $\{(0, 0, 1)\}$ *forma uma base de* $E_T(-1)$.

Exemplo 1.23. *Seja* $T_\alpha : \mathbb{R}^2 \mapsto \mathbb{R}^2$ *a rotação de um vetor* (a, b) *com origem em* $(0, 0)$ *por um ângulo* α. *Se* $\alpha = \pi$, *todos os elementos do* $\mathbb{R}^2$ *são autovetores de autovalor* -1. *Se* $\alpha = 2\pi$, *todos os elementos do* $\mathbb{R}^2$ *são autovetores de autovalor* 1. *Já para outros valores de* α, T_α *não possui nenhum autovetor.*

Seja $S_\alpha : \mathbb{C} \mapsto \mathbb{C}$ *a rotação de um elemento do plano* $\mathbb{C}$ *em relação à origem. Se* $z \in \mathbb{C}$ *pode ser escrito como* $z = a + bi$, *onde* i *é a unidade imaginária, então* $a = |z|cos(\theta)$ *e* $b = |z|sin(\theta)$, *onde* $|z| = \sqrt{a^2 + b^2}$ *e* θ *é o ângulo em relação ao eixo* x *cuja orientação é anti-horária. Portanto*

$$\begin{aligned} z &= |z| \cdot (cos(\theta) + isin(\theta)) \\ &= |z|e^{i\theta} \end{aligned}$$

Então:

$$
\begin{aligned}
S(z) &= S\left(|z|e^{i\theta}\right) \\
&= |z|e^{i(\theta+\alpha)} \\
&= |z|e^{i\theta+i\alpha} \\
&= |z|e^{i\alpha}e^{i\theta} \\
&= e^{i\alpha}\cdot z
\end{aligned}
$$

Ao contrário da transformação linear T, a sua homóloga para os complexos possui autovetores para qualquer valor de α. A transformação S possui autovalor $e^{i\alpha}$ e autovetor qualquer $z \neq 0$ em $\mathbb{C}$.

Exemplo 1.24 (O operador derivada)**.** *Seja V o espaço das funções reais de uma variável infinitamente deriváveis em um intervalo I. Seja $D : V \mapsto V$ a transformação linear que leva uma função f a sua derivada f'. Uma função f satisfazendo a equação*

$$D(f) = \lambda f$$

com $\lambda \neq 0$. É equivalente a dizer que f é solução da equação diferencial

$$\frac{d}{dx}f(x) = \lambda f(x)$$

Fazendo $f(x) = e^{\lambda g(x)}$ e se limitando ao domínio em que $f(x) > 0$, obtemos

$$\lambda \cdot g'(x)e^{\lambda g(x)} = \lambda e^{\lambda g(x)}$$

Sabendo que $\lambda \neq 0$ e que a exponencial nunca é zero, a equação se reduz a

$$g'(x) = 1 \implies g(x) = x + C$$

onde C é uma constante qualquer. E então

$$f(x) = e^{\lambda\cdot(x+C)}$$

Como $f(x) > 0$ é satisfeito para qualquer real x e qualquer constante C, então f é solução da equação diferencial em todo o domínio $\mathbb{R}$. Isso significa que qualquer função $e^{\lambda\cdot(x+C)}$ é um autovetor com autovalor λ em relação ao operador derivada.

No caso $\lambda = 0$, temos a equação diferencial

$$\frac{d}{dx}f(x) = 0$$

onde 0 é a função identicamente nula. A solução (e portanto o autovetor) é $f(x) = C$, onde C é uma constante real qualquer.

Exemplo 1.25 (O operador integral)**.** *Seja V o espaço das funções contínuas em um intervalo I. Se $f \in V$, definamos $g(x)$ da seguinte forma:*

$$g(x) = \int_a^x f(t)\,\mathrm{d}t$$

Para encontrar um autovetor do operador integral, usamos a equação (66)*:*

$$\int_a^x f(t)\,\mathrm{d}t = \lambda f(x)$$

$$\frac{d}{dx}\int_a^x f(t)\,\mathrm{d}t = \lambda \frac{d}{dx} f(x)$$

Pelo teorema fundamental do cálculo:

$$f(x) = \lambda f'(x)$$

e caímos na mesma equação diferencial do exemplo (1.24)*, mas com autovalor $\frac{1}{\lambda}$. Ou seja, $f(x) = e^{\frac{1}{\lambda}(x+C)}$ se $\lambda \neq 0$.*

Exemplo 1.26 (Subespaços invariantes)**.** *Seja $T : V \mapsto W$ uma transformação linear que admita a existência de autovetores de autovalor λ. Seja $x \in E(\lambda)$, $x \neq 0$. Portanto se $y = cx$,*

$$T(y) = T(cx) = cT(x) = c\lambda x = \lambda cx = \lambda y$$

então y também é um autovetor de autovalor λ.

Denomina-se que um subespaço U é invariante sob a transformação S se S é uma operação fechada em U. Portanto $L(\{x\})$ é invariante sob a transformação T.

Teorema 1.45. *Se o conjunto $U = \{u_1, \ldots, u_k\}$ é formado por k autovetores de uma transformação linear $T : V \mapsto W$ correspondendo a k autovalores distintos, então U é linearmente independente.*

Demonstração. Se $k = 1$, U é independente. Suponha que para um certo $k = p - 1$ a afirmação seja verdadeira. Suponha que para um certo p, existem escalares a_i tais que

$$\sum_{i=0}^{p} a_i u_i = 0 \tag{67}$$

Aplicando T na equação, obtemos

$$T\left(\sum_{i=0}^{p} a_i u_i\right) = T(0)$$
$$\sum_{i=0}^{p} a_i T(u_i) = 0$$
$$\sum_{i=0}^{p} a_i \lambda_i u_i = 0 \tag{68}$$

Subtraindo (68) de equação (67) vezes λ_p, obtemos

$$\sum_{i=0}^{p} a_i \lambda_i u_i - a_i \lambda_p u_i = 0$$
$$\sum_{i=0}^{p-1} a_i (\lambda_i - \lambda_p) u_i = 0$$

Por hipótese, os primeiros $p-1$ autovetores são linearmente independentes com autovalores distintos. Então $a_i(\lambda_i - \lambda_p) = 0$. O valor $\lambda_i - \lambda_p$ não pode ser zero porque os autovalores são diferentes. Portanto $\forall i,\ a_i = 0$ e $\lambda_i \neq \lambda_p$, o que significa que a afirmação também é verdadeira para p. Pelo princípio da indução, a afirmação é válida para todo k. ■

Teorema 1.46 (Bases de autovetores e diagonalização)**.** *Seja V um espaço com $\dim V = n$ e $T : V \mapsto V$ uma transformação linear. Então T só pode ter até no máximo n autovalores distintos. Se T tem exatamente n autovalores distintos, então é possível construir uma base ξ de V com os distintos autovetores de T. Além disso, a matriz que representa T_ξ^ξ é uma matriz diagonal cujos elementos não-nulos são os autovalores de T.*

Demonstração. Se T tivesse mais do que n autovalores distintos, pelo teorema (1.45) haveria mais de n autovetores distintos, o que implicaria no absurdo de haver um conjunto com mais de n elementos de V independentes entre si, o que contrariaria $\dim V = n$. Se T tem exatamente n autovalores distintos, pelo teorema (1.45) haveria n autovetores linearmente independentes, no que esses poderiam formar uma base ξ de V.

Pelo teorema (1.44), a base ξ garante que T_ξ^ξ é uma matriz diagonal e, como demonstrado nesse mesmo teorema, as suas colunas são

$$[T]_\xi^\xi = (\lambda_1 I_1, \lambda_2 I_2, \dots, \lambda_n I_n)$$

e portanto os elementos da diagonal são os autovalores associados aos autovetores de ξ. ■

Note que a ordem dos escalares (autovalores) da matriz diagonal $[T]_\xi^\xi$ depende da ordem dos elementos (autovetores) de ξ.

Exemplo 1.27. *Seja $V = C(0,\pi)$ o espaço vetorial real de todas as funções reais contínuas no intervalo $[0,\pi]$. Seja S o subespaço de todas as funções f que possuem derivadas de segunda ordem contínuas e que verificam as condições de fronteira $f(0) = f(\pi) = 0$. Seja $T : S \mapsto V$ a transformação linear que leva uma função f à sua segunda derivada f''. Então*

$$T(f) = \lambda f \implies f'' = \lambda f$$

Fazendo $f(x) = e^{\alpha x}$, obtemos

$$\alpha^2 e^{\alpha x} = \lambda e^{\alpha x}$$
$$\alpha = \pm\sqrt{\lambda}$$

Então $e^{\alpha_1 x}$ e $e^{-\alpha_2 x}$ são soluções da equação diferencial em todo o $\mathbb{R}$, onde $\alpha_1 = +\sqrt{\lambda}$ e $\alpha_2 = -\sqrt{\lambda}$. Como o operador derivada é linear, então a combinação linear dessas duas soluções independentes também é uma solução. E então

$$f(x) = Ae^{\alpha_1 x} + Be^{-\alpha_2 x} \tag{69}$$

com A e B reais arbitrários. Mas como $f \in S$, então f precisa satisfazer as condições $f(0) = f(\pi) = 0$. Suponha que $\lambda > 0$. Fazendo $f(0) = 0$:

$$0 = A + B \implies B = -A$$

e $f(\pi) = 0$:

$$0 = Ae^{\alpha_1 \pi} - Ae^{-\alpha_2 \pi} \implies A = 0 \implies B = 0$$

A solução trivial $f(x) = 0$ não é um autovetor. Agora suponha que $\lambda = 0$. Não podemos usar a equação (69) pois ela somente retornará uma única solução, restando outra. Voltando à equação diferencial $f'' = \lambda f$, duas soluções são os polinômios $f_1(x) = x$ e $f_2(x) = 1$. Fazendo a combinação linear delas, encontramos $f(x) = Ax + B$, que é uma solução mais geral. Testando os valores de contorno:

$$0 = f(0) = A0 + B \implies B = 0$$
$$0 = f(\pi) = A\pi \implies A = 0$$

A possibilidade $\lambda = 0$ somente retorna a solução trivial. Fazendo $\lambda < 0$, podemos usar a equação (69) e

$$\alpha_1 = \sqrt{-|\lambda|} = i\sqrt{\lambda^+} = i\alpha^+$$
$$\alpha_2 = -\sqrt{-|\lambda|} = -i\sqrt{\lambda^+} = -i\alpha^+$$

$$f(x) = Ae^{i\alpha^+ + x} + Be^{-i\alpha^+ x}$$

onde $\lambda^+ = |\lambda|$ *e* $\alpha^+ = \sqrt{\lambda^+}$. *A função* f *precisa ser real, ou seja,* $\overline{f} = f$, *no que implica*

$$\overline{A}e^{-i\alpha^+ x} + \overline{B}e^{i\alpha^+ x} = Ae^{i\alpha^+ x} + Be^{-i\alpha^+ x}$$

$$\overline{A} = B \quad e \quad \overline{B} = A$$

Seja $A = a + bi$ *e* $B = a - bi$. *Então*

$$\begin{aligned} f(x) &= (a + bi)e^{i\alpha^+ x} + (a - bi)e^{-i\alpha^+ x} \\ &= ae^{i\alpha^+ x} + ae^{-i\alpha^+ x} + bie^{i\alpha^+ x} - bie^{-i\alpha^+ x} \\ &= 2a\frac{1}{2}\left[e^{i\alpha^+ x} + e^{-i\alpha^+ x}\right] + 2bi^2\frac{1}{2i}\left[e^{i\alpha^+ x} - e^{-i\alpha^+ x}\right] \\ &= 2a\cos(\alpha^+ x) - 2b\sin(\alpha^+ x) \end{aligned}$$

Usando os valores de contorno:

$$f(0) = 0 \implies 2a = 0 \implies a = 0$$
$$f(\pi) = 0 \implies -2b\sin(\alpha^+ \pi) = 0$$

Não queremos $b = 0$, *pois isso culminaria na solução trivial. Então*

$$\sin(\alpha^+ \pi) = 0$$

e

$$\alpha^+ \pi = n\pi \implies \alpha^+ = n \quad n = 0, 1, 2, 3\ldots$$

$$f(x) = -2b\sin(nx)$$

Portanto os autovetores e os autovalores associados são respectivamente

$$f(x) = c_n \sin(nx) \quad c_n \neq 0$$

e

$$\lambda = \alpha^2 = (\pm i\alpha^+)^2 = -n^2$$

1.8 Operadores em espaços euclidianos e seus autovalores

Teorema 1.47. *Dado V um espaço euclidiano, A um subespaço de V, uma transformação linear $T : A \mapsto V$ que admite um autovalor λ e x um autovetor associado à λ, então*

$$\lambda = \frac{\langle x, T(x) \rangle}{\langle x, x \rangle}$$

Demonstração.

$$\langle x, T(x) \rangle = \langle x, \lambda x \rangle = \lambda \langle x, x \rangle$$

Dado que $x \neq 0$ por ser um autovetor, pelo axioma (14), $\langle x, x \rangle \neq 0$. E então

$$\lambda = \frac{\langle x, T(x) \rangle}{\langle x, x \rangle}$$

■

Se $\lambda \in \mathbb{R}$, então $\langle x, T(x) \rangle = \langle T(x), x \rangle$. Para um $\lambda \in \mathbb{C}$ tal que $\overline{\lambda} = -\lambda$, então $\langle x, T(x) \rangle = - \langle T(x), x \rangle$.

1.8.1 Transformações hermitianas e anti-hermitianas

Definição 1.20. *Seja $T : V \mapsto W$ uma transformação linear, com $V \subseteq W$, onde V é espaço euclidiano. A função T denomina-se hermitiana se*

$$\langle T(x), y \rangle = \langle x, T(y) \rangle \quad \forall x, y \in V$$

Por outro lado, se

$$\langle T(x), y \rangle = - \langle x, T(y) \rangle \quad \forall x, y \in V$$

então T denomina-se anti-hermitiana.

Exemplo 1.28. *Seja $C(a, b)$ o espaço vetorial das funções reais contínuas no intervalo fechado $[a, b]$ com o produto interno definido por*

$$f, g \in C(a, b) \quad \langle f, g \rangle = \int_a^b f(x)g(x)\,\mathrm{d}x$$

Seja também uma transformação linear $T : V \mapsto C(a, b)$, onde V é um subespaço de $C(a, b)$, e $\langle f, T(g) \rangle = \int_a^b f(x)T(g(x))\,\mathrm{d}x$. Se T for hermitiana, então $\langle f, T(g) \rangle = \langle T(f), g \rangle$ e

$$\int_a^b f(x)T(g(x))\,\mathrm{d}x = \int_a^b T(f(x))g(x)\,\mathrm{d}x$$
$$\int_a^b f(x)T(g(x)) - T(f(x))g(x)\,\mathrm{d}x = 0$$

Mas se T for anti-hermitiana, então

$$\int_a^b f(x)T(g(x))\,\mathrm{d}x = -\int_a^b T(f(x))g(x)\,\mathrm{d}x$$
$$\int_a^b f(x)T(g(x)) + T(f(x))g(x)\,\mathrm{d}x = 0$$

Exemplo 1.29 (Multiplicação de uma função por um escalar)**.** *Sejam $C(a,b)$ e T respectivamente o espaço euclidiano e a transformação linear do exemplo* (1.28)*. Definindo a operação de T como $T(f) = cf$, onde $c \in \mathbb{R}$, T adquire a propriedade hermitiana:*

$$\begin{aligned}\langle T(f), g\rangle &= \langle cf, g\rangle \\ &= \int_a^b (cf(x))g(x)\,\mathrm{d}x \\ &= \int_a^b f(x)(cg(x))\,\mathrm{d}x \\ &= \langle f, T(g)\rangle\end{aligned}$$

Exemplo 1.30 (O operador derivada)**.** *Novamente, sejam $C(a,b)$ e T respectivamente o espaço euclidiano e a transformação linear do exemplo* (1.28) *e V como sendo o espaço das funções contínuas f que possuem derivadas contínuas no intervalo fechado $[a,b]$ e satisfazem a condição de fronteira $f(a) = f(b)$. Seja $D : V \mapsto C(a,b),\ D(f) = f'$. Para $f, g \in V$:*

$$\langle f, D(g)\rangle = \int_a^b f(x)g'(x)\,\mathrm{d}x \; \langle D(f), g\rangle = \int_a^b f(x)'g(x)\,\mathrm{d}x$$

Logo

$$\begin{aligned}\langle f, D(g)\rangle + \langle D(f), g\rangle &= \int_a^b f(x)g'(x)\,\mathrm{d}x + \int_a^b f(x)'g(x)\,\mathrm{d}x \\ &= \int_a^b f(x)g'(x) + f(x)'g(x)\,\mathrm{d}x \\ &= \int_a^b \frac{d}{dx}(f(x)g(x))\,\mathrm{d}x \\ &= f(x)g(x)\Big|_a^b \\ &= f(b)g(b) - f(a)g(a) \\ &= 0 \qquad (f(a) = f(b),\ g(a) = g(b))\end{aligned}$$

E então $\langle f, D(g)\rangle = -\langle D(f), g\rangle$, o que mostra que o operador D, no espaço euclidiano descrito, é anti-hermitiano.

Exemplo 1.31 (Operadores de Sturm-Liouville)**.** *Sejam, mais uma vez, $C(a,b)$ e T respectivamente o espaço euclidiano e a transformação linear do exemplo (1.28) e V o subespaço das funções f com derivada de segunda ordem contínuas em $[a,b]$ que satisfazem as condições de fronteira*

$$p(a)f(a) = 0 \qquad p(b)f(b) = 0 \tag{70}$$

onde p é uma função pertencente a $C(a,b)$ com derivada contínua em $[a,b]$. Seja $q \in C(a,b)$. A transformação $T : V \mapsto C(a,b)$ dado pela expressão

$$T(f) = (pf')' + qf \tag{71}$$

é denominado de operador de Sturm-Liouville.

Sejam $f, g \in V$. Note que

$$\begin{aligned}
\langle f, T(g)\rangle &= \int_a^b f(x)(p(x)g(x)')' + f(x)q(x)g(x)\,\mathrm{d}x \\
\langle T(f), g\rangle &= \int_a^b g(x)(p(x)f(x)')' + g(x)q(x)f(x)\,\mathrm{d}x
\end{aligned}$$

Então

$$\langle f, T(g)\rangle - \langle T(f), g\rangle = \int_a^b f(x)(p(x)g(x)')' - \int_a^b g(x)(p(x)f(x)')'\,\mathrm{d}x \tag{72}$$

Usando integração por partes de forma que integre a derivada do produto e derive a função restante, obtemos

$$\begin{aligned}
&\langle f, T(g)\rangle - \langle T(f), g\rangle = \\
&\qquad f(x)p(x)g'(x)\Big|_a^b - \int_a^b f'(x)p(x)g'(x)\,\mathrm{d}x \\
&\qquad\qquad - f'(x)p(x)g(x)\Big|_a^b + \int_a^b f'(x)p(x)g'(x)\,\mathrm{d}x
\end{aligned}$$

$$\begin{aligned}
\langle f, T(g)\rangle - \langle T(f), g\rangle &= (f(x)p(x)g'(x) - f'(x)p(x)g(x))\Big|_a^b \\
&= f(b)p(b)g'(b) - f'(b)p(b)g(b) - f(a)p(a)g'(a) + f'(a)p(a)g(a) \\
&= 0 \qquad \textit{(Condição de fronteira de } f \textit{ e } g\textit{)}
\end{aligned}$$

$$\langle f, T(g)\rangle = \langle T(f), g\rangle$$

Logo o operador de Sturm-Liouville é hermitiano nesse espaço euclidiano.

Os autovetores de T são as funções não-nulas que satisfazem a equação diferencial

$$\lambda f = (pf')' + qf \tag{73}$$

no intervalo $[a, b]$ e com as condições de fronteira de (70).

Teorema 1.48. *Se uma transformação linear T possui λ como um de seus autovalores, então*

- *Se T for hermitiana, então $\overline{\lambda} = \lambda$.*
- *Se T for anti-hermitiana, então $\overline{\lambda} = -\lambda$.*

Demonstração. Seja u um autovetor de T associado a um autovalor λ. Pelo teorema (1.47),

$$\lambda = \frac{\langle u, T(u)\rangle}{\langle u, u\rangle}$$

$$\begin{aligned} \overline{\lambda} &= \frac{\overline{\langle u, T(u)\rangle}}{\langle u, u\rangle} \\ &= \frac{\langle T(u), u\rangle}{\langle u, u\rangle} \end{aligned}$$

Se T for hermitiana, então $\langle T(u), u\rangle = \langle u, T(u)\rangle$ e

$$\overline{\lambda} = \lambda$$

Mas se T for anti-hermitiana, então $\langle T(u), u\rangle = -\langle u, T(u)\rangle$ e

$$\overline{\lambda} = -\lambda$$

■

Se T age em um espaço euclidiano real, então no caso em que for hermitiana o teorema (1.48) não nos fornece nada útil. Mas se T for anti-hermitiana, segue que $\lambda = -\lambda$, no que implica $\lambda = 0$.

1.8.2 Ortogonalidade de autovetores

Teorema 1.49 (Ortogonalidade de autovetores)**.** *Se a transformação linear T é hermitiana ou anti-hermitiana e admitir pelo menos dois autovalores distintos λ e μ associados respectivamente aos autovetores x e y, então $\langle x, y \rangle = 0$.*

Demonstração. Pela definição de autovetor:

$$\langle T(x), y \rangle = \langle \lambda x, y \rangle = \overline{\lambda} \langle x, y \rangle$$

Se T for hermitiana, então $\langle T(x), y \rangle = \langle x, T(y) \rangle$ e isso significa que

$$\langle x, T(y) \rangle = \langle x, \mu y \rangle = \mu \langle x, y \rangle$$
$$\overline{\lambda} \langle x, y \rangle = \mu \langle x, y \rangle$$

Pelo teorema (1.48), $\overline{\lambda} = \lambda$, e então

$$\lambda \langle x, y \rangle = \mu \langle x, y \rangle$$

Fazer $\langle x, y \rangle \neq 0$ implica $\lambda = \mu$, o que contradiz a hipótese. Então $\langle x, y \rangle = 0$.

Se T for anti-hermitiana, então $\langle T(x), y \rangle = -\langle x, T(y) \rangle$ e pelo teorema (1.48), $\overline{\lambda} = -\lambda$. Portanto

$$\overline{\lambda} \langle x, y \rangle = -\mu \langle x, y \rangle$$
$$\lambda \langle x, y \rangle = \mu \langle x, y \rangle$$

no que implica $\langle x, y \rangle = 0$. ■

Exemplo 1.32. *Voltemos à equação diferencial* (73) *dos autovalores do operador de Sturm-Liouville:*

$$\lambda f = (pf')' + qf \tag{74}$$

Dado que T é hermitiano, pelo teorema (1.49)*, se essa equação diferencial admitir duas funções autovetores f e g associados respectivamente aos autovalores λ e μ que obedecem as condições de fronteira* (70)*, então f e g serão ortogonais.*

Fazendo $p(x) = 1$, $q(x) = 0$ e $\lambda = -k^2$, com $k \neq 0$, na equação (74)*, obtemos*

$$f'' + k^2 f = 0 \tag{75}$$

que é a equação diferencial do oscilador harmônico simples da mecânica clássica. Todas as suas soluções são dadas por

$$f(t) = A \sin(kt) + B \cos kt$$

Verificando as condições de fronteira (relembrando $p(x) = 1$):

$$f(0) = 0 \implies B = 0$$

$$f(\pi) = 0 \implies A \sin(k\pi) = 0$$

Para f ser um autovetor, $A \neq 0$, no que implica $\sin(k\pi) = 0$. Portanto $k\pi = n\pi \implies k = n$, com n inteiro. Ou seja, T assume autovalores $\lambda = -n^2$ e autovetores $f(t) = c_n \sin(nt)$.

Uma consequência disso é que, pelo teorema (1.49),

$$\int_0^\pi \sin(nt)\sin(mt)\,\mathrm{d}x = 0 \qquad \textit{se} \qquad m^2 \neq n^2$$

Exemplo 1.33. *Seja $C_w(a,b)$ o espaço das funções reais contínuas no intervalo $[a,b]$ com o produto interno definido por*

$$f, g \in C_w(a,b) \qquad \langle f, g\rangle = \int_a^b w(x)f(x)g(x)\,\mathrm{d}x$$

onde $w(x) \in C_w(a,b)$, $w(x) > 0$. O novo operador de Sturm-Liouville nesse espaço é dado por

$$T_w : V \mapsto C_w(a,b)$$
$$T_w(f) = \frac{(pf')' + qf}{w}$$

onde V é o espaço das funções com segunda derivada contínua e obedecem as condições de contorno de (70).

Dadas duas funções $f, g \in V$, verifica-se que

$$\begin{aligned}\langle f, T(g)\rangle &= \int_a^b \frac{1}{w(x)}w(x)f(x)(p(x)g'(x))' + \frac{1}{w(x)}w(x)f(x)q(x)g(x)\,\mathrm{d}x \\ &= \int_a^b f(x)(p(x)g'(x))' + f(x)q(x)g(x)\,\mathrm{d}x\end{aligned}$$

$$\begin{aligned}\langle T(f), g\rangle &= \int_a^b \frac{1}{w(x)}w(x)g(x)(p(x)f'(x))' + \frac{1}{w(x)}w(x)g(x)q(x)f(x)\,\mathrm{d}x \\ &= \int_a^b g(x)(p(x)f'(x))' + g(x)q(x)f(x)\,\mathrm{d}x\end{aligned}$$

E então retornamos à equação (72), *no que implica que T é hermitiana nesse espaço euclidiano.*

Exemplo 1.34. *Seja V um subespaço de um espaço euclidiano complexo W, $T : V \mapsto W$ uma transformação linear e uma função $Q : V \mapsto \mathbb{C}$ de V definida por*

$$Q(x) = \langle T(x), x \rangle \quad \forall x \in V$$

Se T for hermitiana, então

$$Q(x) = \langle T(x), x \rangle = \langle x, T(x) \rangle = \overline{Q(x)} \implies Q(x) \in \mathbb{R}$$

Mas se T for anti-hermitiana:

$$Q(x) = \langle T(x), x \rangle = -\langle x, T(x) \rangle = -\overline{Q(x)} \implies Q(x) \in L(\{i\})$$

Além disso, se t é um escalar complexo, então

$$Q(tx) = \langle T(tx), tx \rangle = t \langle tT(x), x \rangle = |t|^2 \langle tT(x), x \rangle = |t|^2 Q(x)$$

Dados $x, y \in V$ e $t \in \mathbb{C}$:

$$\begin{aligned}
Q(x+ty) &= \langle T(x+ty), x+ty \rangle \\
&= \langle T(x+ty), x \rangle + \langle T(x+ty), ty \rangle \\
&= \langle T(x) + T(ty), x \rangle + t \langle T(x) + T(ty), y \rangle \\
&= \langle T(x), x \rangle + \langle T(ty), x \rangle + t \langle T(x), y \rangle + t \langle T(ty), y \rangle \\
&= Q(x) + \bar{t} \langle T(y), x \rangle + t \langle T(x), y \rangle + |t|^2 Q(y)
\end{aligned}$$

Se $Q(x) = 0,\ \forall x \in V$, então

$$\langle T(x), x \rangle = 0 \implies T(x) = x = 0 \quad \forall x \in V \quad (\textit{Axioma 14})$$

E se $Q(x)$ é real para todo x em V, implica em

$$\overline{Q(x+ty)} = Q(x+ty) \quad \overline{Q(x)} = Q(x) \quad \overline{Q(y)} = Q(y)$$

e então

$$\overline{Q(x)} + t \langle x, T(y) \rangle + \bar{t} \langle y, T(x) \rangle + |t|^2 \overline{Q(y)} = \\ Q(x) + \bar{t} \langle T(y), x \rangle + t \langle T(x), y \rangle + |t|^2 Q(y)$$

$$\begin{aligned}
t \langle x, T(y) \rangle + \bar{t} \langle y, T(x) \rangle &= \bar{t} \langle T(y), x \rangle + t \langle T(x), y \rangle \\
\langle x, T(ty) \rangle + \langle ty, T(x) \rangle &= \langle T(ty), x \rangle + \langle T(x), ty \rangle
\end{aligned}$$

Agora suponha $y = x$ e fixemos $t = 1$. Segue que

$$\begin{aligned}
\langle x, T(x) \rangle + \langle x, T(x) \rangle &= \langle T(x), x \rangle + \langle T(x), x \rangle \\
\langle x, T(x) \rangle &= \langle T(x), x \rangle
\end{aligned}$$

ou seja, T é hermitiana. Isso mostra que T é hermitiana se e somente se $Q(x) \in \mathbb{R},\ \forall x \in V$.

Exemplo 1.35 (Polinômios de Legendre e o operador de Sturm-Liouville)**.** *Seja $P_n(t)$ o polinômio de Legendre de grau n dado por*

$$P_n(t) = \frac{1}{2^n n!} f_n^{(n)}(t) \quad \textit{onde} \quad f_n(t) = (t^2 - 1)^n \tag{76}$$

Derivando o termo $f_n(t)$, obtemos

$$f_n'(t) = n \cdot (2t) \cdot (t^2 - 1)^{n-1}$$

Multiplicando ambos os lados por $(t^2 - 1)$, temos que

$$\begin{aligned}(t^2 - 1) f_n'(t) &= 2nt \cdot (t^2 - 1)^n \\ (t^2 - 1) f_n'(t) &= 2nt f_n(t) \end{aligned} \tag{77}$$

Derivando ambos os lados:

$$2t f_n^{(1)}(t) + (t^2 - 1) f_n^{(2)}(t) = 2n f_n(t) + 2nt f_n^{(1)}(t) \tag{78}$$

Repetindo o processo várias vezes, obtemos as seguintes equações:

$$\begin{aligned} 2f_n^{(1)}(t) + 4t f_n^{(2)}(t) + (t^2 - 1) f_n^{(3)}(t) &= 4n f_n^{(1)}(t) + 2nt f_n^{(2)}(t) \\ 6f_n^{(2)}(t) + 6t f_n^{(3)}(t) + (t^2 - 1) f_n^{(4)}(t) &= 6n f_n^{(2)}(t) + 2nt f_n^{(3)}(t) \\ 12f_n^{(3)}(t) + 8t f_n^{(4)}(t) + (t^2 - 1) f_n^{(5)}(t) &= 8n f_n^{(3)}(t) + 2nt f_n^{(4)}(t) \\ 20f_n^{(4)}(t) + 10t f_n^{(5)}(t) + (t^2 - 1) f_n^{(6)}(t) &= 10n f_n^{(4)}(t) + 2nt f_n^{(5)}(t) \end{aligned}$$

É razoável supor que o múltiplo do primeiro termo é a sequência $a_0 = 0$, $a_n = a_{n-1} + 2n$, onde n é a quantidade de derivadas feitas na equação (77). Então

$$a_{n-1} = a_{n-2} + 2(n-1) \implies a_n = a_{n-2} + 2(n-1) + 2n$$

Repetindo o processo até a_0, podemos escrever

$$a_n = \sum_{i=0}^{n} a_0 + 2i = \sum_{i=0}^{n} 2i = 2\sum_{i=0}^{n} i = 2\frac{n(n+1)}{2} = n(n+1)$$

Para o segundo e quinto termos, aparentemente temos uma progressão aritmética $b_0 = 2$, $b_n = 2(n+1)$. E os demais termos, os múltiplos são os mesmos. Agora suponha que derivando a equação (77) $n-1$ vezes obteremos

$$\begin{aligned}(n-1)n f_n^{(n-1)}(t) + 2nt f_n^{(n)}(t) + (t^2 - 1) f_n^{(n+1)}(t) &= \\ 2n^2 f_n^{(n-1)}(t) + 2nt f_n^{(n)}(t) \end{aligned} \tag{79}$$

Derivando ambos os lados:

$$(n-1)nf_n^{(n)}(t)+2nf_n^{(n)}(t)+2ntf_n^{(n+1)}(t)$$
$$+2tf_n^{(n+1)}(t)+(t^2-1)f_n^{(n+2)}(t)=$$
$$2n^2f_n^{(n)}(t)+2nf_n^{(n)}(t)+2ntf_n^{(n+1)}(t)$$

$$(n-1+2)nf_n^{(n)}(t)+2t(n+1)f_n^{(n+1)}(t)+(t^2-1)f_n^{(n+2)}(t)=$$
$$2n(n+1)f_n^{(n)}(t)+2ntf_n^{(n+1)}(t)$$

$$n(n+1)f_n^{(n)}(t)+2t(n+1)f_n^{(n+1)}(t)+(t^2-1)f_n^{(n+2)}(t)=$$
$$2n(n+1)f_n^{(n)}(t)+2ntf_n^{(n+1)}(t) \quad (80)$$

que é a equação (79) com n derivadas. Portanto, por indução, a n-ésima derivada de (77) é dada pela equação (80).

Pela equação (76):

$$f_n^{(n)}(t)=2^n n!\cdot P_n(t)$$

Substituindo em (80):

$$n(n+1)!2^nP_n(t)+2^{n+1}(n+1)!tP_n'+2^nn!\cdot(t^2-1)P_n''=$$
$$2^{n+1}n\cdot(n+1)!P_n(t)+2^{n+1}n\cdot n!tP_n'$$

Dividindo ambos os lados por $2^nn!$:

$$n(n+1)P_n(t)+2t(n+1)P_n'+(t^2-1)P_n''=2n(n+1)P_n(t)+2ntP_n'$$

Isolando o termo $P_n(t)$:

$$n(n+1)(1-2)P_n(t)=2ntP_n'-2t(n+1)P_n'-(t^2-1)P_n''$$
$$-n(n+1)P_n(t)=(2tn-2t(n+1))P_n'-(t^2-1)P_n''$$
$$n(n+1)P_n(t)=2tP_n'+(t^2-1)P_n''$$

A última expressão pode ser reescrita como

$$((t^2-1)P_n')'=n(n+1)P_n(t) \quad (81)$$

Comparando com a expressão (74):

$$(pf')' + qf = \lambda f$$

a equação diferencial (81) *é a do autovalor* $\lambda = n(n+1)$ *do operador de Sturm-Liouville com* $p(t) = t^2 - 1$ *e* $q(t) = 0$. *Então* $P_n(t)$ *é o autovetor associado ao autovalor* $\lambda = n(n+1)$. *A condição de fronteira* $p(a)P_n(a) = p(b)P_n(b) = 0$ *é satisfeita quando* a *e* b *são raízes de* $p(x) = t^2 - 1 = 0$, *ou seja,* $a = -1$ *e* $b = 1$. *O intervalo trabalhado é, portanto,* $[-1, 1]$.

Teorema 1.50 (Bases ortogonais de autovetores de transformações hermitianas)**.** *Dado um espaço euclidiano* V *de dimensão finita* $\dim V = n$*, se uma transformação* $T : V \mapsto V$ *for hermitiana (ou anti-hermitiana), então existe uma base ortogonal de autovetores de* T*.*

Demonstração. Se $n = 1$, então dada uma base $\{e_1\}$, $x = a_1 e_1$, $a_1 \neq 0$, e $T(x) = a_2 e_1$, temos

$$T(x) = a_2 \cdot \frac{a_1}{a_1} e_1 = \frac{a_2}{a_1} x$$

e portanto x é um autovetor de autovalor $\frac{a_2}{a_1}$ que constitui uma base ortogonal de V.

Agora suponha que o teorema (1.50) seja válido para qualquer espaço euclidiano de dimensão $n-1$. Dado um autovetor u_1 associado ao autovalor λ_1, esse elemento pode fazer parte de uma base de V, $B = \{u_1, v_2, \ldots, v_n\}$, que é ortogonal (o método de ortogonalização de Gram-Schmidt garante a existência dessa base). Agora seja V' o subespaço de V formado pelos elementos x que são ortogonais a u_1, ou seja, $\langle x, u_1 \rangle = 0$. Por x pertencer também ao espaço V, podemos escrever x como uma combinação linear da base B:

$$x = a_1 u_1 + a_2 v_2 + a_3 v_3 + \ldots a_n v_n$$

Então $\langle x, u_1 \rangle = a_1 \langle u_1, u_1 \rangle$. O coeficiente a_1 é zero pela hipótese de x pertencer ao subespaço dos elementos ortogonais a u_1. Portanto

$$x = a_2 v_2 + a_3 v_3 + \ldots a_n v_n$$

o que implica que o conjunto $B' = \{v_2, \ldots, v_n\}$ é uma base de V' e portanto $\dim V' = n-1$. Agora dada uma transformação linear hermitiana T, temos que

$$\langle T(x), u_1 \rangle = \langle x, T(u_1) \rangle = \langle x, \lambda_1 u_1 \rangle = \lambda_1 \langle x, u_1 \rangle = 0$$

Isso significa que se x for ortogonal a u_1, então sua imagem em T, $T(x)$, também é ortogonal a u_1. Ou seja, o espaço euclidiano V' é fechado em relação à transformação hermitiana T. Então podemos usar a hipótese que foi dada no início para afirmar que existe uma base do subespaço V' de autovetores $B'' = \{u_2, u_3, \ldots, u_n\}$ de T. Incluindo u_1 à base B'' obtemos uma outra base de V com n autovetores de T ortogonais entre si, o que completa a demonstração por indução. ∎

1.8.3 Representação matricial dos operadores hermitianos e anti-hermitianos

Teorema 1.51. *Seja V um espaço euclidiano complexo de dimensão finita e $T : V \mapsto V$ uma transformação linear. Se $\{e_1, \ldots, e_n\}$ for uma base de V, então*

- *T é hermitiana se e somente se $\langle T(e_i), e_j\rangle = \langle e_i, T(e_j)\rangle$ para todo par i, j.*
- *T é anti-hermitiana se e somente se $\langle T(e_i), e_j\rangle = -\langle e_i, T(e_j)\rangle$ para todo par i, j.*

Demonstração. Se T é hermitiana ou anti-hermitiana, o resultado é trivial.

Agora suponha que $\langle T(e_i), e_j\rangle = \langle e_i, T(e_j)\rangle$ para todo par i, j. Seja $x, y \in V$. Dado que $\{e_1, \ldots, e_n\}$ é uma base de V, então x e y podem ser escritos da seguinte forma:

$$x = \sum_{i=1}^{n} a_i e_i \qquad y = \sum_{j=1}^{n} b_j e_j$$

Calculando, então, o produto interno $\langle T(x), y\rangle$:

$$\begin{aligned}\langle T(x), y\rangle &= \left\langle \sum_{i=1}^{n} a_i T(e_i), \sum_{j=1}^{n} b_j e_j \right\rangle \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} \overline{a_i} b_j \langle T(e_i), e_j\rangle\end{aligned}$$

Por hipótese, $\langle T(e_i), e_j\rangle = \langle e_i, T(e_j)\rangle$. Logo:

$$\begin{aligned}\langle T(x), y\rangle &= \sum_{i=1}^{n}\sum_{j=1}^{n} \overline{a_i} b_j \langle e_i, T(e_j)\rangle \\ &= \left\langle \sum_{i=1}^{n} a_i e_i, \sum_{j=1}^{n} b_j T(e_j) \right\rangle \\ &= \langle x, T(y)\rangle\end{aligned}$$

Para T anti-hermitiana, o processo é idêntico mas com a diferença que $\langle T(e_i), e_j\rangle = -\langle e_i, T(e_j)\rangle$ para qualquer i e j.

■

Teorema 1.52. *Se $\beta = \{e_1, \ldots, e_n\}$ é uma base finita ortonormal de V, $T : V \mapsto V$ uma transformação linear e $A = (a_{ij})$ a representação matricial de T_β^β, então*

- *T é hermitiana se somente se $\overline{a_{ij}} = a_{ij}$ para quaisquer i e j.*

- *T é anti-hermitiana se somente se $\overline{a_{ij}} = -a_{ij}$ para quaisquer i e j.*

Demonstração. Sabemos que é possível escrever $[T]_\beta^\beta$ da seguinte forma:

$$[T]_\beta^\beta = (T_\beta^\beta(I_1), \dots, T_\beta^\beta(I_n))$$

onde I_i é o elemento e_i na base β no formato de matriz coluna. Como β é uma base ortonormal, vimos que qualquer produto interno em V pode ser escrito como

$$\langle x, y\rangle = \sum_{k=0}^{n} \overline{a_i} b_i$$

onde a_i e b_i são respectivamente as coordenadas de x e y na base β. Semelhantemente podemos reescrever a equação como

$$\langle \overline{x}, y\rangle = \sum_{k=0}^{n} a_i b_i$$

Dada a matriz $A = (a_{ij}) = [T]_\beta^\beta$, o elemento a_{ij} da linha i e coluna j de $[T]_\beta^\beta$ pode ser calculado pela expressão

$$a_{ij} = \langle \overline{I_i}, T(I_j)\rangle = \langle I_i, T(I_j)\rangle$$

onde escolhemos uma coluna j e emparelhamo-os com I_i, no que resultará na soma de $n-1$ termos nulos mais $1 \cdot a_{ij}$.

Se T for hermitiana, $a_{ij} = \langle I_i, T(I_j)\rangle = \langle T(I_j), I_i\rangle = \overline{a_{ij}}$ para qualquer i e j. Procedimento análogo para T anti-hermitiana.

Agora se tratando da recíproca: se $\overline{a_{ij}} = a_{ij}$ para quaisquer i e j, isso também significa que $\langle T(I_j), I_i\rangle = \langle I_j, T(I_i)\rangle$ para quaisquer i e j. Sabendo que I_i e I_j são a forma matricial dos elementos da base β na coordenada relativa a β, então, pelo teorema (1.51), T é hermitiana. Argumento semelhante para T anti-hermitiana. ■

1.8.4 Matrizes hermitianas, unitárias e ortogonais

Definição 1.21. *Se $A = (a_{ij})$ é uma matriz quadrada tal que $\overline{a_{ij}} = a_{ji}$ para qualquer i e j, então denominamos A de matriz hermitiana. Analogamente, chamamos A de matriz anti-hermitiana se $\overline{a_{ij}} = -a_{ji}$ para todo i e j. Além disso, a matriz conjugada de A é dado por $\overline{A} = (\overline{a_{ij}})$.*

Seja $T : V \mapsto V$ uma transformação linear hermitiana no espaço euclidiano complexo V e β uma base ortonormal. Então dados $x, y \in V$, o produto interno entre eles pode ser visto como

$$\langle x, y\rangle = Y^T \cdot \overline{X}$$

onde X e Y são as matrizes colunas respectivas de x e y na base β e Y^T é a transposta de Y. Seja $A = [T]_\beta^\beta$ Então

$$\langle T(x), x\rangle = \langle x, T(x)\rangle \implies X^T \cdot \overline{T(X)} = T(X)^T \cdot \overline{X}$$

$$X^T \cdot \overline{AX} = X^T (A)^T \overline{X}$$

$$\begin{aligned} X^T \cdot \overline{AX} - X^T (A)^T \overline{X} &= 0 \\ X^T (\overline{A} - A^T)\overline{X} &= 0 \end{aligned}$$

Como o resultado é sempre zero para qualquer $x \in V$ escolhido, segue que $\overline{A} - A^T = 0$. E então podemos enunciar o seguinte teorema:

Teorema 1.53. *Se um espaço euclidiano V possuir uma base ortonormal β, $T : V \mapsto V$ for uma transformação linear hermitiana e A é a matriz que representa T na base β, então*

$$\overline{A} = A^T \quad \textit{ou ainda} \quad A = \overline{A^T}$$

Analogamente, se T for anti-hermitiana:

$$A = -\overline{A^T}$$

Definição 1.22. *Dada uma matriz A, denominamos a matriz $A^* = \overline{A^T}$ de a transconjugada de A.*

Em outras palavras, o teorema 1.53 diz que se T é hermitiana, então $A = A^*$ (ou equivalentemente, $[T]_\beta^\beta$ é hermitiana).

Definição 1.23. *Uma matriz quadrada A se diz unitária se $AA^* = I$. Se diz ortogonal se $AA^T = I$. A matriz A se diz simétrica se $A = A^T$, ou anti-simétrica se $A = -A^T$.*

Exemplo 1.36. *Seja $A = (a_{ij})$ uma matriz unitária. Se A só contém elementos reais, ou seja, $\overline{a_{ij}} = a_{ij}$, então A também é ortogonal, visto que $A^* = A^T$.*

Exemplo 1.37 (Determinante de uma matriz ortogonal)**.** *Se A é uma matriz ortogonal, então* $\det(AA^T) = \det(I)$ *e*

$$\begin{aligned} \det(A)\det(A^T) &= 1 \\ \det(A)^2 &= 1 \quad (\det(A^T) = \det(A)) \\ \det(A) &= \pm 1 \end{aligned}$$

ou seja, ou $\det A = 1$ *ou* $\det A = -1$.

1.8.5 Transformações unitárias

Definição 1.24. *Seja V um subespaço de W, um espaço euclidiano. Uma transformação linear $T : V \mapsto W$ se diz unitária quando*

$$\langle T(x), T(y) \rangle = \langle x, y \rangle \quad \forall x, y \in V \tag{82}$$

Se W for um espaço real e T obedecer (82), então T também pode ser denominado de transformação ortogonal.

Teorema 1.54. *Se $T : V \mapsto W$ é uma transformação unitária, então para $x, y \in V$ é válido:*

1. $\langle x, y \rangle = 0 \implies \langle T(x), T(y) \rangle = 0$
2. $\|T(x)\| = \|x\|$
3. $\|T(x) - T(y)\| = \|x - y\|$
4. *T possui uma inversa e T^{-1} é unitária em $T(V)$.*

Demonstração. 1. O resultado é trivial, uma vez que

$$0 = \langle x, y \rangle = \langle T(x), T(y) \rangle$$

2. $\|x\| = \sqrt{\langle x, x \rangle} = \sqrt{\langle T(x), T(x) \rangle} = \|T(x)\|$
3. $\|T(x) - T(y)\| = \|T(x - y)\| = \|x - y\|$
4. Para mostrarmos que T possui uma inversa, basta provar que $N(T) = \{0\}$. Seja x um elemento de V tal que $T(x) = 0$. Então

$$\langle T(x), T(x) \rangle = \langle 0, 0 \rangle = 0$$

Mas $\langle T(x), T(x) \rangle = \langle x, x \rangle$ e portanto, pelo axioma (14), $x = 0$. E então, pelo teorema (1.26), T é biunívoca e possui uma inversa T^{-1}.

Agora seja $y_1, y_2 \in T(V)$. Então

$$\begin{aligned} \langle y_1, y_2 \rangle &= \langle T(T^{-1}(y_1)), T(T^{-1}(y_2)) \rangle \\ &= \langle T^{-1}(y_1), T^{-1}(y_2) \rangle \end{aligned}$$

o que implica que a transformação T^{-1} também é unitária.

■

Teorema 1.55. *Seja $T : V \mapsto W$ uma transformação unitária. Então*

1. *Se T possui algum autovalor λ, então $|\lambda| = 1$.*

2. *Se x e y são autovetores associados aos autovalores distintos λ e σ, então x e y são ortogonais.*

Demonstração. 1. Se λ é um autovalor da transformação unitária T e x um autovetor associado a λ, então

$$\langle T(x), T(x)\rangle = \langle \lambda x, \lambda x\rangle = \lambda^2 \|x\|^2$$

Por outro lado,

$$\langle T(x), T(x)\rangle = \langle x, x\rangle = \|x\|^2$$

Portanto $\lambda^2\|x\|^2 = \|x\|^2$. Como x é um autovetor, $\|x\| \neq 0$ e então $\lambda^2 = 1$, ou $|\lambda| = 1$

2. Se $T(x) = \lambda x$ e $T(y) = \sigma y$, $\lambda \neq \sigma$, então

$$\langle T(x), T(y)\rangle = \langle \lambda x, \sigma y\rangle = \sigma\overline{\lambda}\,\langle x, y\rangle$$

Por T ser unitária, temos também $\langle T(x), T(y)\rangle = \langle x, y\rangle$, no que implica

$$\sigma\overline{\lambda}\,\langle x, y\rangle = \langle x, y\rangle \tag{83}$$

A igualdade só pode ser satisfeita se $\langle x, y\rangle = 0$ ou $\sigma\overline{\lambda} = 1$. Pelo item 1 do teorema (1.55), $|\sigma| = |\lambda| = 1$. Então supomos

$$\lambda = e^{i\theta} \quad \sigma = e^{i\phi}$$

A equação $\sigma\overline{\lambda} = 1$ nos diz que

$$e^{i\phi}e^{-i\theta} = e^{i(\phi-\theta)} = 1 \implies \phi = \theta \implies \sigma = \lambda$$

o que é falso. Portanto a equação (83) só pode ser verdade se $\langle x, y\rangle = 0$.

■

Teorema 1.56. *Seja V um espaço euclidiano com $\dim V = n$ e $\beta = \{e_1, \ldots, e_n\}$ uma base de V. Uma transformação linear $T : V \mapsto V$ é unitária em V se e somente se*

$$\langle T(e_i), T(e_j)\rangle = \langle e_i, e_j\rangle$$

para quaisquer i e j.

Demonstração. Seja $x, y \in V$. Então podemos escrever x e y na base β:

$$x = \sum_{i=1}^{n} a_i e_i \qquad y = \sum_{j=1}^{n} b_j e_j$$

então

$$\begin{aligned}\langle x, y \rangle &= \left\langle \sum_{i=1}^{n} a_i e_i, \sum_{j=1}^{n} b_j e_j \right\rangle \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} \overline{a_i} b_j \langle e_i, e_j \rangle\end{aligned}$$

Se $\langle T(e_i), T(e_j) \rangle = \langle e_i, e_j \rangle$ para quaisquer i e j, então

$$\begin{aligned}\langle x, y \rangle &= \sum_{i=1}^{n} \sum_{j=1}^{n} \overline{a_i} b_j \langle T(e_i), T(e_j) \rangle \\ &= \langle T(x), T(y) \rangle\end{aligned}$$

Já a recíproca, se T for unitária, trivialmente $\langle T(e_i), T(e_j) \rangle = \langle e_i, e_j \rangle$ para quaisquer i e j. ■

Teorema 1.57. *Se V é um espaço euclidiano de dimensão finita n e $\beta = \{e_1, \ldots, e_n\}$ é uma base ortonormal de V, então uma transformação linear $T : V \mapsto V$ é unitária se e somente se $[T]_\beta^\beta$ for uma matriz unitária, ou seja,*

$$([T]_\beta^\beta)^* [T]_\beta^\beta = I$$

Demonstração. A matriz $[T]_\beta^\beta$ pode ser escrita como

$$[T]_\beta^\beta = (T_\beta^\beta(I_1), \ldots, T_\beta^\beta(I_n))$$

onde I_i é a forma matricial do elemento e_i na base β. Além disso, o produto interno pode ser dado por

$$x, y \in V \quad \langle x, y \rangle = \sum_{i=1}^{n} \overline{a_i} b_i$$

Essa equação também pode ser reescrita como

$$x, y \in V \quad \langle \overline{x}, y \rangle = \sum_{i=1}^{n} a_i b_i \tag{84}$$

Suponha que $([T]_\beta^\beta)^* [T]_\beta^\beta = I$.

Note que, pela multiplicação de matrizes e usando a equação (84),

$$\langle T_\beta^\beta(I_i), T_\beta^\beta(I_j)\rangle = \delta_{ij}$$

onde δ_{ij} é o elemento da linha i e coluna j da matriz identidade de ordem n. Isso também é válido para $\langle I_i, I_j\rangle$, uma vez que β é ortonormal. Então podemos fazer a igualdade $\langle T_\beta^\beta(I_i), T_\beta^\beta(I_j)\rangle = \langle I_i, I_j\rangle$ para quaisquer i e j. Pelo teorema (1.56), T é unitária.

Agora vamos à recíproca. Suponha que a transformação T é unitária e que $([T]_\beta^\beta)^*[T]_\beta^\beta = A$, onde $A = (a_{ij})$ é uma matriz quadrada. Usando multiplicação de matrizes:

$$a_{ij} = \langle T(I_i), T(I_j)\rangle$$

Mas T é unitária, então

$$a_{ij} = \langle I_i, I_j\rangle = \delta_{ij}$$

Logo $([T]_\beta^\beta)^*[T]_\beta^\beta = I$ e, portanto, a matriz $[T]_\beta^\beta$ é unitária. ■

Teorema 1.58. *Se A é uma matriz unitária, então*

1. *A possui uma matriz inversa e $A^{-1} = A^*$*
2. *As matrizes A^T, $\overline{A}$ e A^* também são unitárias.*
3. *Os autovalores de A são complexos de norma igual a um.*
4. *$|\det A| = 1$, e se A possuir apenas elementos reais, $\det A = \pm 1$.*

Demonstração.

1. A equação $A^*A = I$ implica $\det(A^*A) = \det I$ e então

$$\det(A^*)\det A = 1$$

o que significa que $\det A \neq 0$ e portanto possui uma inversa. Então $A^*A = I$ nos diz que A^* é a inversa à esquerda de A. Como a inversa é única, então $AA^* = I$ e $A^{-1} = A^*$.

2. Aplicando a transposta em $A^*A = I$ obtemos $A^T\overline{A} = I$, o que significa que $\overline{A}$ é unitária. Além disso, A^T é a inversa de $\overline{A}$, o que significa que $\overline{A}A^T = I$. Portanto a matriz A^T é unitária. Similarmente, $AA^* = I$, o que significa que A^* também é uma matriz unitária.

3. Se A é unitária, então, pelo teorema (1.57), está associada à alguma transformação unitária T. Por sua vez, se T tem autovalores (que são os mesmos de A), pelo teorema (1.55), esses terão módulo 1.

4. A expressão $A^*A = I$ implica em

$$\begin{aligned}
\det(A^*A) &= \det I \\
\det(A^*) \det A &= 1 \\
\det\left(\overline{A^T}\right) \det A &= 1 \\
\overline{\det(A^T)} \det A &= 1 \\
\overline{\det A} \det A &= 1 \\
|\det A|^2 &= 1 \\
|\det A| &= 1
\end{aligned}$$

Se A for real, os únicos números reais que possuem módulo um é 1 e -1. Portanto, $\det A = \pm 1$.

■

2 Revisão de Cálculo

2.1 Convergência de sequências e séries

2.1.1 Sequências

Definição 2.1 (Sequência e limite)**.** *Uma lista de termos ordenados $\{a_n\}$, onde n é um inteiro positivo e a_n uma função cujo domínio são os valores de n, é denominada de sequência numérica, ou sucessão.*

Se para qualquer $\epsilon > 0$ existir algum inteiro positivo N tal que

$$|a_n - L| < \epsilon \quad \forall n \geq N$$

então é dito que a sequência $\{a_n\}$ possui limite L, é convergente e denotado como

$$\lim_{n \mapsto \infty} a_n = L \tag{85}$$

Se uma sequência não for convergente, então dizemos que ela é divergente.

Os valores a_n e L podem ser reais ou complexos. Se este último for o caso, podemos fazer $a_n = u + vi$ e $L = a + bi$ e então podemos repartir $a_n - L$ em uma parte real e uma imaginária:

$$a_n - L = u(n) - a + (v(n) - b)i$$

Os termos $|u(n) - a|$, $|v(n) - b|$ e $|a_n - L|$ são os comprimentos de um triângulo retângulo no plano $\mathbb{C}$ e, portanto, são válidas as desigualdades:

$$\begin{cases} |u(n) - a| \leq |a_n - L| \\ |v(n) - b| \leq |a_n - L| \end{cases} \tag{86}$$

$$|a_n - L| \leq |u(n) - a| + |v(n) - b| \tag{87}$$

O conjunto de desigualdades (86) nos diz que se a_n tende a L quando n tende a infinito, pelo teorema do confronto $u(n)$ tenderá a a e $v(n)$ tenderá a b.

Já na inequação (87), se $u(n)$ tende a a e $v(n)$ tende a b, então a_n tende a L. Isso significa que se uma sequência de números complexos convergir, então ambas as sequências associadas à parte real e à parte imaginária vão convergir, e a recíproca também é verdadeira, no que resulta

$$\lim_{n \to \infty} a_n = \lim_{n \to \infty} u(n) + i \lim_{n \to \infty} v(n)$$

Exemplo 2.1. *Seja $x \in \mathbb{R}$, $|x| < 1$ e a sequência $\{a_n = x^n\}$. Dado um $\epsilon > 0$, seja $N = \lceil \log_{|x|} \epsilon \rceil$, denotando o inteiro superior mais próximo ou igual a $\log_{|x|} \epsilon$. Então $\left|x^N\right| < \epsilon$. Fazendo $L = 0$, a expressão $|a_n - L|$ se reduz a $|x^n|$. Além disso, multiplicando $|x| < 1$ por $|x|$ obtemos*

$$0 < |x|^2 < |x| < 1 \tag{88}$$

Repetindo o processo mais N vezes obtemos a inequação $\left|x^{N+1}\right| < \left|x^N\right|$. Dessa forma, descobrimos que $|x^n| \leq \left|x^N\right| < \epsilon$ para todo $n \geq N$, o que implica

$$\lim_{n\to\infty} x^n = 0$$

Definição 2.2 (Divergências que tendem ao infinito)**.** *Uma sequência $\{a_n\}$ é dita que diverge para mais infinito se*

$$\forall y \quad \exists N \text{ tal que } a_n > y \quad \forall n \geq N$$

e escrevemos como $\lim\limits_{n\to\infty} a_n = +\infty$.

Analogamente, $\{a_n\}$ diverge para menos infinito se

$$\forall y \quad \exists N \text{ tal que } a_n < y \quad \forall n \geq N$$

e é denotado como $\lim\limits_{n\to\infty} a_n = -\infty$.

Para sequências de números complexos, cujo corpo não possui ordem, diz-se que a_n tende a ∞ quando $n \mapsto \infty$ se $|a_n| \mapsto \infty$.

Teorema 2.1 (Combinação linear de sequências)**.** *Se $\{a_n\}$ e $\{b_n\}$ são duas sequências convergentes e $\alpha, \beta \in \mathbb{C}$, então $\{\alpha a_n + \beta b_n\}$ converge e*

$$\alpha \lim_{n\to\infty} a_n + \beta \lim_{n\to\infty} b_n = \lim_{n\to\infty} \alpha a_n + \beta b_n$$

Demonstração. Dado $\epsilon_1 > 0$ e $\epsilon_2 > 0$, é verdade que existem N_1 e N_2 tal que

$$|a_n - L_1| < \epsilon_1 \quad \forall n \geq N_1 \quad \text{e} \quad |b_n - L_2| < \epsilon_2 \quad \forall n \geq N_2$$

Multiplicando a primeira inequação por $|\alpha|$ e a segunda por $|\beta|$:

$$|\alpha a_n - \alpha L_1| < |\alpha|\epsilon_1 \quad \forall n \geq N_1 \quad \text{e} \quad |\beta b_n - \beta L_2| < |\beta|\epsilon_2 \quad \forall n \geq N_2$$

Fazendo $N = \max\{N_1, N_2\}$ e somando ambas as expressões, obtemos

$$|\alpha a_n - \alpha L_1| + |\beta b_n - \beta L_2| < |\alpha|\epsilon_1 + |\beta|\epsilon_2 \quad \forall n \geq N$$

Mas pela desigualdade triangular:

$$|\alpha a_n - \alpha L_1 + \beta b_n - \beta L_2| \leq |\alpha a_n - \alpha L_1| + |\beta b_n - \beta L_2|$$

e portanto, fazendo $\epsilon = |\alpha|\epsilon_1 + |\beta|\epsilon_2$:

$$|(\alpha a_n + \beta b_n) - (\alpha L_1 + \beta L_2)| < \epsilon \quad \forall n \geq N$$

o que implica que $\{\alpha a_n + \beta b_n\}$ converge e

$$\begin{aligned}\lim_{n\to\infty} \alpha a_n + \beta b_n &= \alpha L_1 + \beta L_2 \\ &= \alpha \lim_{n\to\infty} a_n + \beta \lim_{n\to\infty} b_n\end{aligned}$$

■

2.1.2 Séries

Definição 2.3. *Dada a sequência* $\{a_n\}$*, a soma parcial* s_n *dos termos dessa sequência é dada por*

$$s_n = \sum_{k=0}^{n} a_k$$

A sequência definida por $\{s_n\}$ *é chamado de série infinita. E se*

$$\lim_{n\to\infty} s_n = S$$

onde S *é um número real ou complexo, então dizemos que* $\{s_n\}$ *converge para* S*.*

A série infinita $\{s_n\}$ advinda da sequência $\{a_n\}$ também pode ser escrita na forma

$$\lim_{n\to\infty} \sum_{k=0}^{n} a_n$$

que denotamos com uma expressão mais simples:

$$\sum_{k=0}^{\infty} a_n$$

Se $\{s_n\}$ diverge, então $\sum_{k=0}^{\infty} a_k$ diverge e não há soma. É importante notar que a soma é o resultado da série $\sum_{k=0}^{\infty} a_k$ quando esta converge. O termo $\sum_{k=0}^{\infty} a_k$ diz respeito à série infinita, podendo ela convergir ou divergir.

Teorema 2.2. *Se* $\sum a_k$ *e* $\sum b_k$ *são duas séries convergentes e* $\alpha, \beta \in \mathbb{C}$*, então* $\sum \alpha a_k + \beta b_k$ *também converge e*

$$\alpha \sum_{k=0}^{\infty} a_n + \beta \sum_{k=0}^{\infty} b_n = \sum_{k=0}^{\infty} \alpha a_k + \beta b_k$$

Demonstração. Sabendo que $\sum_{k=0}^{\infty} a_n = \lim_{n\to\infty} \sum_{k=0}^{n} a_n$ e o mesmo para b_n, a demonstração é trivial usando o teorema (2.1). ∎

O subíndice do somatório não necessariamente precisa começar com $k = 0$, pois isso não afeta na convergência ou divergência da série infinita. Se $\sum_{k=0}^{\infty} a_n$ converge, então $\sum_{k=p}^{\infty} a_n$ com $p \geq 0$ também converge, uma vez que

$$\sum_{k=p}^{\infty} a_n = \sum_{k=0}^{\infty} a_n - \sum_{k=0}^{p-1} a_n$$

Sabendo que $\sum_{k=0}^{\infty} a_n$ possui um resultado definido e que $\sum_{k=0}^{p-1} a_n$ é uma soma finita, então $\sum_{k=p}^{\infty} a_n$ é a soma de dois números e logo é convergente. A recíproca também é verdadeira: se $\sum_{k=p}^{\infty} a_n$ convergir, então $\sum_{k=0}^{\infty} a_n$ também converge.

Pela contrapositiva, se $\sum_{k=p}^{\infty} a_n$ diverge, então $\sum_{k=0}^{\infty} a_n$ diverge e vice-versa.

Isso implica que a adição ou subtração de finitos termos não altera a convergência ou divergência de uma série infinita. O teorema a seguir mostra que isso nem sempre é verdade quando adicionamos ou subtraímos outras séries infinitas.

Teorema 2.3. *Se* $\sum_{k=0}^{\infty} a_n$ *converge e* $\sum_{k=0}^{\infty} b_n$ *diverge, então* $\sum_{k=0}^{\infty} a_n + b_n$ *diverge.*

Demonstração.

$$b_n = (a_n + b_n) - a_n$$

e portanto

$$\sum_{k=0}^{\infty} b_n = \sum_{k=0}^{\infty} (a_n + b_n) - a_n$$

Se $\sum_{k=0}^{\infty} a_n + b_n$ converge, então, pelo teorema (2.2):

$$\sum_{k=0}^{\infty} (a_n + b_n) - \sum_{k=0}^{\infty} a_n = \sum_{k=0}^{\infty} (a_n + b_n) - a_n$$

o que leva a conclusão que $\sum_{k=0}^{\infty} (a_n + b_n) - a_n$ é a soma de duas séries convergentes e portanto converge, o que é falso. Portanto $\sum_{k=0}^{\infty} (a_n + b_n)$ não pode convergir e então diverge. ■

Teorema 2.4 (Série telescópica)**.** *Sejam* $\{a_n\}$ *e* $\{b_n\}$ *duas sequências tais que*

$$\forall n \quad a_n = b_{n+1} - b_n$$

Então a série $\sum a_n$ *converge se e somente se* $\{b_n\}$ *convergir e em caso positivo, é válido*

$$\sum_{k=0}^{\infty} a_n = L - b_0 \qquad \textit{onde } L = \lim_{n\to\infty} b_n$$

Demonstração. Note que

$$\begin{aligned} s_0 &= b_1 - b_0 \\ s_1 &= b_2 - b_1 + b_1 - b_0 \\ &= b_2 - b_0 \end{aligned}$$

Suponha que $s_{n-1} = b_n - b_0$ seja válido para um certo n. Demonstra-se por indução que essa expressão é válida para todo n:

$$\begin{aligned} s_{n-1} + (b_{n+1} - b_n) &= b_n - b_0 + (b_{n+1} - b_n) \\ s_{n-1} + a_n &= b_{n+1} - b_0 \\ s_n &= b_{n+1} - b_0 \end{aligned}$$

E logo

$$\sum_{k=0}^{\infty} a_n = \lim_{n\to\infty} b_{n+1} - b_0$$

Se $\{b_n\}$ converge e $\lim_{n\to\infty} b_n = L$, então

$$\sum_{k=0}^{\infty} a_n = L - b_0$$

E se $\sum_{k=0}^{\infty} a_n$ convergir, então

$$\lim_{n\to\infty} b_n = b_0 + \sum_{k=0}^{\infty} a_n$$

o que implica que $\{b_n\}$ também converge.

■

Teorema 2.5 (Série geométrica)**.** *Se* $x \in \mathbb{C}$, $|x| < 1$, *então a série infinita* $\sum x^k$ *converge e*

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x} \tag{89}$$

Demonstração. A soma parcial de $\{x^n\}$ é dada por

$$s_n = 1 + \ldots + x^n$$

Note que se $x \neq 0$, então

$$\begin{aligned}
\frac{s_n}{x} &= \frac{1}{x} + 1 + x + \ldots + x^{n-1} \\
\frac{s_n}{x} - \frac{1}{x} + x^n &= 1 + x + \ldots + x^{n-1} + x^n \\
\frac{s_n - 1}{x} + x^n &= s_n \\
s_n - 1 + x^{n+1} &= x \cdot s_n \\
x^{n+1} - 1 &= s_n(x-1) \\
s_n &= \frac{x^{n+1} - 1}{x - 1} \\
s_n &= \frac{1 - x^{n+1}}{1 - x} \qquad (90)
\end{aligned}$$

E então

$$\sum_{k=0}^{\infty} x^k = \lim_{n\to\infty} s_n = \lim_{n\to\infty} \frac{1 - x^{n+1}}{1 - x}$$

Sabemos que pelo exemplo (2.1), $\lim_{n\to\infty} x^{n+1} = 0$ se $|x| < 1$ e portanto

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1 - x}$$

■

Se $|x| > 1$, então o limite $\lim_{n\to\infty} x^{n+1}$ diverge e não há soma.

Se $|x| = 1$, então $x = e^{i\theta}$ e

$$\lim_{n\to\infty} \frac{1 - x^{n+1}}{1 - x} = \lim_{n\to\infty} \frac{1 - e^{i\theta(n+1)}}{1 - e^{i\theta}}$$

O valor $\theta = 0$ não pode ser possível por causa do denominador. Se $x = 1$, a soma parcial será $s_n = n + 1$, que diverge quando $n \to \infty$. E para $\theta \neq 0$, o valor oscilará entre os complexos de módulo um, sem assumir um valor definido quando $n \to \infty$.

O valor x em (89) pode ser substituído por qualquer função $f(x)$ cujo domínio fique limitado àqueles elementos x que satisfazem $|f(x)| < 1$. Em particular, se $f(x) = -x$, então obtemos

$$\sum_{k=0}^{\infty} (-x)^k = 1 - x + x^2 - x^3 + \ldots = \frac{1}{1 + x} \qquad \text{se } |x| < 1$$

2.1.3 Critérios de convergência

Na prática, nem sempre é fácil obter uma fórmula para as somas parciais s_n e assim verificar se uma série converge ou diverge com base no limite de s_n. Entretanto é possível estabelecer condições suficientes para uma convergência, condições necessárias caso uma série convirja e, finalmente, condições necessárias e suficientes para uma convergência sem precisar encontrar s_n.

Teorema 2.6 (Condição necessária para convergência)**.** *Se $\sum a_n$ converge, então é condição necessária*

$$\lim_{n\to\infty} a_n = 0$$

Demonstração. Seja s_n a soma parcial de a_n. Para um n finito temos

$$s_n = a_1 + a_2 + \ldots + a_n$$

Isolando o termo a_n:

$$a_n = s_n - s_{n-1}$$

Aplicando o limite em ambos os lados:

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} s_n - s_{n-1}$$

Uma vez que $\sum a_n$ converge, então ambos s_n e s_{n-1} também convergem e para um mesmo valor, no que implica $\lim_{n\to\infty} a_n = 0$

■

O teorema (2.6) é mais comumente usado na sua contrapositiva: se $\lim_{n\to\infty} a_n \neq 0$ ou diverge, então $\sum a_n$ diverge.

Teorema 2.7 (Teste de convergência em séries de termos não-negativos)**.** *Se $\{a_n\}$ é uma sequência com $a_n \geq 0$, $\forall n \geq 0$, então $\sum a_n$ converge se e somente se a sua respectiva soma parcial s_n é limitada superiormente.*

Demonstração. Dado que os elementos de $\{a_n\}$ são não-negativos, então $0 \leq s_{n-1} \leq s_n \quad \forall n$. Se s_n é limitada superiormente por M, então $0 \leq s_n \leq M$. Aplicando o limite:

$$0 \leq \lim_{n\to\infty} s_n \leq M$$

o que significa que $\sum a_n$ converge.

E se $\sum a_n$ converge, então s_n é limitada superiormente por $\lim_{n\to\infty} s_n$. ■

Exemplo 2.2. *A série $s = \sum \frac{1}{n!}$ é limitada superiormente. Para mostrar isso, note que enquanto a potência 2^{n-1} só possui 2 como fator, o fatorial de qualquer inteiro $n \geq 2$ inclui 2 e outros inteiros maiores como fatores, no que resulta*

$$n! \geq 2^{n-1}$$

$$\frac{1}{n!} \leq \frac{1}{2^{n-1}} \quad \forall n \geq 1$$

E sabemos que

$$\begin{aligned}
\sum_{n=0}^{\infty} \frac{1}{2^{n-1}} &= \sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^{n-1} \\
&= \sum_{n=0}^{\infty} 2\left(\frac{1}{2}\right)^{n} \\
&= 2 \cdot \frac{1}{1-\frac{1}{2}} \\
&= 4
\end{aligned}$$

E portanto

$$\sum_{n=0}^{\infty} \frac{1}{n!} \leq 4$$

Pelo teorema (2.7)*, a série* $\displaystyle\sum_{n=0}^{\infty} \frac{1}{n!}$ *converge.*

Teorema 2.8 (Teste de convergência usando integral)**.** *Se f é uma função tal que $f'(x) < 0$, $f(x) > 0$ e é integrável para todo $x \geq 1$ e*

$$s_n = \sum_{k=1}^{n} f(k) \quad t_n = \int_1^n f(x)\,\mathrm{d}x$$

para todo $n \geq 1$, então ambos s_n e t_n convergem ou ambos divergem.

Demonstração. A somatória $\displaystyle\sum_{k=2}^{n} f(k)$ diz respeito à integral da função escada inferior a f enquanto que $\displaystyle\sum_{k=1}^{n-1} f(k)$ é a integral da função escada superior a f. Uma vez que f se situa entre essas duas funções escada, então

$$\sum_{k=2}^{n} f(k) \leq \int_1^n f(x)\,\mathrm{d}x \leq \sum_{k=1}^{n-1} f(k)$$

No limite em que $n \to \infty$, se $\displaystyle\int_1^n f(x)\,\mathrm{d}x$ divergir, então $\displaystyle\sum_{k=1}^{n-1} f(k)$ diverge. Se $\displaystyle\sum_{k=2}^{n} f(k)$ diverge, então $\displaystyle\int_1^n f(x)$ diverge. Se $\displaystyle\sum_{k=1}^{n-1} f(k)$ converge, então $\displaystyle\int_1^n f(x)\,\mathrm{d}x$ converge. Finalmente, se $\displaystyle\int_1^n f(x)\,\mathrm{d}x$ converge, então $\displaystyle\sum_{k=2}^{n} f(k)$ converge. ∎

Exemplo 2.3. *Seja a série* $\sum_{n=1}^{\infty} \frac{1}{n^s}$*. Sabemos que* $f(x) = \frac{1}{x^s}$ *é monótona decrescente no intervalo* $[1, \infty)$ *se* $s > 0$*. Para* $s \neq 1$*:*

$$\begin{aligned} \int_1^n \frac{\mathrm{d}x}{x^s} &= \left(-\frac{1}{s \cdot x^{s-1}} \right) \Bigg|_1^n \\ &= \frac{1}{s} - \frac{1}{s \cdot n^{s-1}} \end{aligned}$$

$$\lim_{n\to\infty} \int_1^n \frac{\mathrm{d}x}{x^s} = \lim_{n\to\infty} \frac{1}{s} - \frac{1}{s \cdot n^{s-1}} \tag{91}$$

Se $s < 1$*, então a potência de* n^{s-1} *é negativa, no que implica*

$$\lim_{n\to\infty} \frac{1}{s} - \frac{1}{s \cdot n^{s-1}} = \lim_{n\to\infty} \frac{1}{s} - \frac{n^{1-s}}{s}$$

onde $1 - s > 0$*. Consequentemente, o limite da integral diverge e portanto* $\sum_{n=1}^{\infty} \frac{1}{n^s}$ *diverge se* $s < 1$*.*

Se $s > 1$*, o limite* (91) *converge para* $\frac{1}{s}$*, o que implica que* $\sum_{n=1}^{\infty} \frac{1}{n^s}$ *converge.*

Se $s = 1$*, então*

$$\begin{aligned} \int_1^n \frac{\mathrm{d}x}{x} &= \ln(x) \Big|_1^n \\ &= \ln(n) \end{aligned}$$

E então $\lim_{n\to\infty} \int_1^n \frac{\mathrm{d}x}{x}$ *diverge, o que também deve acontecer para* $\sum_{n=1}^{\infty} \frac{1}{n}$

Teorema 2.9 (Teste da comparação)**.** *Sejam os termos* $a_n \geq 0$ *e* $b_n \geq 0$*,* $\forall n \geq 1$*. Se existir uma constante positiva* c *tal que*

$$a_n \leq c \cdot b_n \quad \forall n \tag{92}$$

então $\sum a_n$ *converge se* $\sum b_n$ *convergir.*

Demonstração. Suponha que s_n e t_n são as somas parciais de a_n e b_n respectivamente. Se a relação (92) for válida para alguma constante positiva c, então se somarmos (92) múltiplas vezes com $n = 1, 2, 3, \ldots, k$ obtemos

$$s_n \leq c \cdot t_n \quad \forall n$$

Fazendo o limite $n \to \infty$:

$$\sum_{n=1}^{\infty} a_n \leq c \sum_{n=1}^{\infty} b_n$$

Se $\sum_{n=1}^{\infty} b_n$ convergir, então $\sum_{n=1}^{\infty} a_n$ é limitado superiormente e então, pelo teorema (2.7), $\sum_{n=1}^{\infty} a_n$ converge. ■

A contrapositiva do teorema (2.9) nos diz que se (92) é satisfeito e $\sum a_n$ diverge, então $\sum b_n$ também diverge.

Além disso, remover os primeiros termos da série não afeta a convergência ou divergência, o que implica que podemos usar o teorema (2.9) com (92) sendo satisfeito para todo $n \geq N$, sendo N um inteiro positivo arbitrário.

Teorema 2.10 (Critério da raiz em séries de termos não-negativos)**.** *O critério da raiz (ou de Cauchy) afirma que se $\sum a_n$ é uma série com termos não-negativos tais que*

$$\lim_{n \to \infty} \sqrt[n]{a_n} = R \tag{93}$$

então a série converge se $R < 1$ e diverge se $R > 1$. E se $R = 1$, o resultado é inconclusivo.

Demonstração. Começamos pela desigualdade

$$0 \leq a_n \leq x^n \quad \forall n \tag{94}$$

Pelo teste da comparação, $\sum a_n$ converge se $\sum x^n$ convergir, o que só é possível se $0 < x < 1$ (descartamos $x = 0$, uma vez que a_n nem sempre é nulo). A desigualdade (94) também pode ser escrita como

$$0 \leq \sqrt[n]{a_n} \leq x$$

Aplicando o limite $n \to \infty$ nessa desigualdade resulta em

$$0 \leq \lim_{n \to \infty} \sqrt[n]{a_n} \leq x$$
$$0 \leq R \leq x$$

Se $R < 1$, então podemos escolher um x tal que $R < x < 1$ e aplicar o teste da comparação, ou seja, $\sum a_n$ converge.

Se $R > 1$, então haverão infinitos termos $\sqrt[k]{a_k} > 1$, o que também significa que haverão infinitos termos $a_k > 1$ e então $\lim_{n \to \infty} a_n$ não pode convergir para 0, no que $\sum a_n$ diverge pelo teorema (2.6).

Agora vejamos dois exemplos de séries infinitas tais que $R = 1$: para $\sum_{n=0}^{\infty} \left(1 + \frac{1}{n}\right)^n$, temos $\lim_{n\to\infty} \sqrt[n]{a_n} = \lim_{n\to\infty} 1 + \frac{1}{n} = 1$, mas

$$\lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n = e \neq 0$$

e portanto a série diverge. Para $\sum_{n=0}^{\infty} \frac{1}{n^2}$, temos $\lim_{n\to\infty} \sqrt[n]{a_n} = \lim_{n\to\infty} \sqrt[n]{\frac{1}{n}} = 1$. Mas sabemos que pelo exemplo (2.3), $\sum_{n=0}^{\infty} \frac{1}{n^2}$ converge. ■

Teorema 2.11 (Critério do quociente em séries de termos não-negativos)**.** *O critério do quociente (ou de D'Alembert) afirma que se $\sum a_n$ é uma série com termos positivos tais que*

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = L \tag{95}$$

então a série converge se $L < 1$ e diverge se $L > 1$. E se $L = 1$, o resultado é inconclusivo.

Demonstração. Suponha $L < 1$. Então existe algum real x tal que $L < x < 1$. Como

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} < x$$

então existe algum inteiro N tal que $\frac{a_{n+1}}{a_n} < x \quad \forall n \geq N$. Manipulando essa desigualdade, obtemos:

$$\begin{aligned} a_{n+1} &< x \cdot a_n \\ \frac{a_{n+1}}{x^{n+1}} &< \frac{a_n}{x^n} \quad \forall n \geq N \end{aligned}$$

o que significa que a sequência $\frac{a_n}{x^n}$ é monotonamente decrescente a partir de $n \geq N$. Ou seja, $\frac{a_n}{x^n} \leq \frac{a_N}{x^N} \quad \forall n \geq N$. Multiplicando ambos os lados por x^n, que é positivo, resulta em

$$a_n \leq \left(\frac{a_N}{x^N}\right) x^n$$

Uma vez que $\frac{a_N}{x^N}$ é positivo e que $x < 1 \implies \sum x^n$ converge, pelo teorema (2.9), $\sum a_n$ também converge.

Se $L > 1$, segue que a_n é sempre crescente a partir de algum N, no que $\lim_{n\to\infty} a_n \neq 0$ e $\sum a_n$ diverge.

No caso $L = 1$, podemos encontrar $\lim_{n\to\infty} \dfrac{\left(1+\dfrac{1}{n+1}\right)^{n+1}}{\left(1+\dfrac{1}{n}\right)^{n}} = 1$, sendo que $\displaystyle\sum_{n=0}^{\infty}\left(1+\frac{1}{n}\right)^{n}$ diverge. E também $\lim_{n\to\infty} \dfrac{\frac{1}{(n+1)^2}}{\frac{1}{n^2}} = \lim_{n\to\infty} \dfrac{n^2}{(n+1)^2} = 1$, sabendo que $\displaystyle\sum_{n=1}^{\infty}\frac{1}{n^2}$ converge. ■

Definição 2.4 (Séries alternadas)**.** *Uma série $\sum a_n$ é dita alternada se pode ser escrita na forma*

$$\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty}(-1)^{n-1}b_n$$

com $b_n > 0 \quad \forall n \geq 1$.

Teorema 2.12 (Convergência em séries alternadas: regra de Leibniz)**.** *Se a sequência $\{a_n\}$ é sempre monótona e decrescente e satisfazer*

$$\lim_{n\to\infty} a_n = 0$$

então a série alternada $\displaystyle\sum_{n=1}^{\infty}(-1)^{n-1}a_n$ converge. Se essa série possui soma S e s_n é a sua soma parcial, então também são válidas as desigualdades

$$0 < (-1)^n(S - s_n) \leq a_{n+1} \quad \forall n \geq 1 \tag{96}$$

Demonstração. Seja s_n a soma parcial de $\displaystyle\sum_{n=1}^{\infty}(-1)^{n-1}a_n$. Os termos pares da série possuem $(-1)^{n-1} = -1$. Portanto a sequência $\{s_{2n}\}$ é monotonamente crescente, uma vez que $s_{2n+2} - s_{2n} = a_{2n+1} - a_{2n+2}$, que é sempre positivo, dado que por hipótese $\{a_n\}$ é monotonamente decrescente. Analogamente, os termos ímpares possuem $(-1)^{n-1} = 1$ e portanto $s_{2n+1} - s_{2n-1} = a_{2n+1} - a_{2n} < 0$, o que implica que $\{s_{2n-1}\}$ é monotonamente decrescente. Sabendo que

$$s_{2n-1} - s_{2n} = a_{2n-1} > 0 \implies s_{2n-1} > s_{2n} \quad \forall n$$

então podemos afirmar que ambas as sequências s_{2n} e s_{2n-1} são limitadas inferiormente pelo primeiro termo de s_{2n}, que é s_2, e limitadas superiormente pelo primeiro termo de s_{2n-1}, que é s_1. Isso significa que ambas convergem quando $n \to \infty$ e ainda

$$\lim_{n\to\infty} s_{2n-1} - s_{2n} = \lim_{n\to\infty} a_{2n-1} = 0$$

e então

$$\lim_{n\to\infty} s_{2n-1} = \lim_{n\to\infty} s_{2n} = S \implies \lim_{n\to\infty} s_n = S$$

Portanto $\sum_{n=1}^{\infty}(-1)^{n-1}a_n$ converge e é igual a S.

Para provar (96), sabemos que s_{2n-1} é monotonamente decrescente e s_{2n} é monotonamente crescente e portanto

$$S \leq s_{2n+1} < s_{2n-1} \qquad s_{2n} < s_{2n+2} \leq S \quad \forall n$$

Juntando ambas de maneira conveniente e multiplicando a primeira por -1:

$$-s_{2n} \geq -S \geq -s_{2n+1} > -s_{2n-1} \qquad s_{2n} < s_{2n+2} \leq S \leq s_{2n+1} \quad \forall n$$

Somando s_{2n-1} à primeira desigualdade e subtraindo s_{2n} da segunda obtemos

$$s_{2n-1} - s_{2n} \geq s_{2n-1} - S \geq s_{2n-1} - s_{2n+1} > 0$$
$$0 < s_{2n+2} - s_{2n} \leq S - s_{2n} \leq s_{2n+1} - s_{2n} \quad \forall n$$

Omitindo os terceiro e segundo termos respectivamente e calculando $s_{2n-1} - s_{2n}$ e $s_{2n+1} - s_{2n}$ resulta

$$0 < s_{2n-1} - S \leq a_{2n} \qquad 0 < S - s_{2n} \leq a_{2n+1} \quad \forall n$$

Sabendo que $2n-1$ é ímpar e que $2n$ é par, então podemos reescrever essas desigualdades da seguinte forma:

$$0 < (-1)^n(S - s_n) \leq a_{n+1}$$

o que completa a demonstração. ■

Exemplo 2.4. *Sabemos que* $\sum_{n=1}^{\infty}\frac{1}{n}$ *diverge, como visto em* (2.3). *Mas como* $\left\{\frac{1}{n}\right\}$ *é sempre monotonamente decrescente e satisfaz* $\lim_{n\to\infty}\frac{1}{n} = 0$, *então pela regra de Leibniz, a sua versão alternada* $\sum_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n}$ *converge.*

Teorema 2.13 (Convergência condicional e absoluta)**.** *Se* $\sum |a_n|$ *converge, então* $\sum a_n$ *também converge e*

$$\left|\sum_{n=1}^{\infty} a_n\right| \leq \sum_{n=1}^{\infty} |a_n| \qquad (97)$$

Demonstração. Primeiramente suponha que os termos de $\{a_n\}$ são todos reais. Fazendo $b_n = a_n + |a_n|$, podemos chegar na desigualdade $0 \leq b_n \leq 2|a_n|$. Por hipótese, $\sum |a_n|$ converge e, portanto, $\sum b_n$ é limitado superiormente, no que implica que $\sum b_n$ converge e então $\sum a_n = \sum b_n - |a_n|$ também converge.

Agora suponha que os termos de $\{a_n\}$ em geral são complexos. Fazendo $a_n = u_n + i \cdot v_n$, com $u_n, v_n \in \mathbb{R}$, podemos chegar na desigualdade $|u_n| \leq |a_n|$. Como $\sum |a_n|$ converge, então $\sum |u_n|$ é limitado superiormente e portanto converge. Como mostrado no parágrafo anterior, a convergência de $\sum |u_n|$ implica na convergência de $\sum u_n$. Da desigualdade $|v_n| \leq |a_n|$, também conclui-se que $\sum |v_n|$ converge e portanto $\sum v_n$ também. A combinação linear de séries convergentes é uma série convergente, então $\sum a_n = \sum u_n + i \cdot v_n$ converge.

Para provar (97), basta notar a desigualdade triangular generalizada

$$\left|\sum_{n=1}^{k} a_n\right| \leq \sum_{n=1}^{k} |a_n|$$

Tomando o limite $k \to \infty$, obtemos (97). ■

Definição 2.5 (Convergência absoluta)**.** *Uma série* $\sum a_n$ *é dita que converge absolutamente se* $\sum |a_n|$ *converge. E se* $\sum a_n$ *converge e* $\sum |a_n|$ *diverge, então* $\sum a_n$ *é semi-convergente.*

Dada duas séries $\sum a_n$ e $\sum b_n$ que convergem absolutamente, então a combinação linear das duas séries também converge absolutamente, uma vez que pela desigualdade triangular

$$\sum_{n=1}^{k} |\alpha a_n + \beta b_n| \leq |\alpha| \sum_{n=1}^{k} |a_n| + |\beta| \sum_{n=1}^{k} |b_n| \leq |\alpha| \sum_{n=1}^{\infty} |a_n| + |\beta| \sum_{n=1}^{\infty} |b_n|$$

$\sum |\alpha a_n + \beta b_n|$ é limitado superiormente e então converge.

Teorema 2.14 (Fórmula de Abel para Soma parcial)**.** *Sejam* $\{a_n\}$ *e* $\{b_n\}$ *duas sequências de números complexos e*

$$A_n = \sum_{k=1}^{n} a_k \tag{98}$$

então é válido a seguinte identidade:

$$\sum_{k=1}^{n} a_k b_k = A_n b_{n+1} + \sum_{k=1}^{n} A_k (b_k - b_{k+1}) \tag{99}$$

Demonstração. Definamos $A_0 = 0$. De (98) podemos obter $a_k = A_k - A_{k-1}$, e portanto

$$\sum_{k=1}^{n} a_k b_k = \sum_{k=1}^{n} (A_k - A_{k-1}) b_k = \sum_{k=1}^{n} A_k b_k - \sum_{k=1}^{n} A_{k-1} b_k \tag{100}$$

mas dado que $A_0 = 0$, podemos fazer

$$\sum_{k=1}^{n} A_{k-1}b_k = \sum_{k=2}^{n} A_{k-1}b_k = \sum_{k=1}^{n-1} A_k b_{k+1} = \left(\sum_{k=1}^{n} A_k b_{k+1}\right) - A_n b_{n+1}$$

e substituindo em (100) resulta

$$\sum_{k=1}^{n} a_k b_k = \sum_{k=1}^{n} A_k b_k - \left(\left(\sum_{k=1}^{n} A_k b_{k+1}\right) - A_n b_{n+1}\right)$$

e finalmente

$$\sum_{k=1}^{n} a_k b_k = A_n b_{n+1} + \sum_{k=1}^{n} A_k (b_k - b_{k+1})$$

■

Fazendo o limite $n \to \infty$ em (99), podemos concluir que se a série $\sum A_k(b_k - b_{k+1}$ convergir e também a sequência $\{A_n b_{n+1}\}$, então $\sum a_k b_k$ converge.

Teorema 2.15 (Teste de Dirichlet)**.** *Seja $\sum a_n$ uma série de elementos complexos cuja soma parcial é uma sequência limitada. Se $\{b_n\}$ é uma sequência que é decrescente e $\lim_{n\to\infty} b_n = 0$, então a série $\sum a_n b_n$ converge.*

Demonstração. Seja $A_n = s_n = \sum_{k=1}^{n} a_k$ do teorema (2.14). Como s_n é limitado superiormente, então existe algum $M > 0$ tal que $\forall n,\ |A_n| \leq M$. Portanto o limite $\lim_{n\to\infty} A_n b_{n+1}$ continua sendo zero. Então para provarmos que $\sum a_k b_k$ converge, precisamos somente mostrar que nas condições da hipótese $\sum A_k(b_k - b_{k+1})$ converge. Como $\{b_n\}$ é decrescente, então $b_k - b_{k+1} > 0 \quad \forall k$ e podemos multiplicar a desigualdade $|A_n| \leq M$ por $b_k - b_{k+1}$:

$$|A_k(b_k - b_{k+1})| \leq M(b_k - b_{k+1})$$

A série $\sum M(b_k - b_{k+1})$ é telescópica e converge, dado que $\lim_{n\to\infty} b_n$ converge. Então $\sum |A_k(b_k - b_{k+1})|$ é limitada superiormente pela série telescópica, o que implica que $\sum A_k(b_k - b_{k+1})$ converge absolutamente e portanto, pelo teorema (2.13), $\sum A_k(b_k - b_{k+1})$ também converge, o que completa a demonstração. ■

Teorema 2.16 (Teste de Abel)**.** *Se $\sum a_n$ é uma série convergente de termos complexos e $\{b_n\}$ é uma sequência monótona convergente de termos reais, então a série $\sum a_n b_n$ converge.*

Demonstração. Seja A_n o termo definido no teorema (2.14). Se $\sum a_n$ converge, então a sequência $\{A_n\}$ converge e consequentemente é limitada. Como $\{b_n\}$ é uma sequência convergente, então $\{A_n b_{n+1}\}$ também converge. E então retornamos à demonstração de (2.15). Se $\{b_n\}$ é monotonamente crescente, usamos o termo $b_{k+1} - b_k$ ao invés de $b_k - b_{k+1}$ para multiplicar $|A_n| \leq M$. A série $\sum M(b_{k+1} - b_k)$ também é telescópica e converge. ■

Teorema 2.17. *Se $\theta \in \mathbb{R}$ não é um múltiplo inteiro de π, então*

$$\sum_{k=1}^{n} e^{2ik\theta} = \frac{\sin(n\theta)}{\sin(\theta)} e^{i(n+1)\theta} \tag{101}$$

e também

$$\left| \sum_{k=1}^{n} e^{2ik\theta} \right| \leq \frac{1}{|\sin(\theta)|} \tag{102}$$

Demonstração. Fazendo $x = e^{2i\theta}$ na soma parcial geométrica (90), obtemos

$$\sum_{k=1}^{n} e^{2ik\theta} = \frac{1 - e^{2i(n+1)\theta}}{1 - e^{2i\theta}} = \frac{1 - e^{2in\theta}}{1 - e^{2i\theta}} e^{2i\theta} \cdot \frac{e^{-in\theta}}{e^{-in\theta}} \cdot \frac{e^{-i\theta}}{e^{-i\theta}}$$

$$\sum_{k=1}^{n} e^{2ik\theta} = \frac{2i}{2i} \cdot \frac{e^{-in\theta} - e^{in\theta}}{e^{-i\theta} - e^{i\theta}} \cdot e^{2i\theta + in\theta - i\theta} = \frac{-\sin(n\theta)}{-\sin(\theta)} e^{i(n+1)\theta}$$

que então leva a (101). Para provar (102), note que $|\sin(n\theta)| \leq 1$ e que $\left|e^{i(n+1)\theta}\right| = 1$. ■

Exemplo 2.5. *Fazendo $a_n = x^n$ no teste de Dirichlet, onde $|x| = 1$ mas $x^2 \neq 1$, então para qualquer sequência monotonamente decrescente $\{b_n\}$ cujo limite é zero a seguinte série converge:*

$$\sum_{n=1}^{\infty} b_n x^n$$

2.2 Integrais impróprias

Definição 2.6 (Integral imprópria do tipo 1)**.** *Seja $f(x)$ uma função integrável em um intervalo $[a, b]$, onde $b \geq a$. Uma nova função $I(x)$ é definida pela expressão*

$$I(x) = \int_a^x f(t)\,\mathrm{d}t$$

Uma integral imprópria do tipo 1 é definida por $\lim_{x\to\infty} I(x)$ e também denotada por $\int_a^\infty f(t)\,\mathrm{d}t$. É dito que essa integral converge quando o limite $\lim_{x\to\infty} I(x)$ existe e é finito. Se não convergir, diz-se que a integral diverge. Se $\lim_{x\to\infty} I(x) = A$, então A é o valor da integral imprópria e denotamos $\int_a^\infty f(t)\,\mathrm{d}t = A$.

Exemplo 2.6. *Dado $f(x) = x^{-s}$, a sua integral em um intervalo $[1, b]$ é dado por*

$$\int_1^b x^{-s}\,\mathrm{d}x = \begin{cases} \dfrac{b^{1-s}-1}{1-s} & \text{se } s \neq 1 \\ \ln(b) & \text{se } s = 1 \end{cases}$$

Se $s = 1$, a integral $\int_1^\infty x^{-s}\,\mathrm{d}x$ diverge, visto que o logaritmo cresce ilimitadamente. No caso $s \neq 1$, $\int_1^\infty x^{-s}\,\mathrm{d}x$ só converge se $\lim_{b\to\infty} b^{1-s}$ converge, o que só acontece quando $1 - s < 0$, ou equivalentemente $s > 1$. Nessa condição, o limite $\lim_{b\to\infty} b^{1-s}$ é zero e portanto

$$\int_1^\infty x^{-s}\,\mathrm{d}x = \frac{1}{s-1}$$

Exemplo 2.7 (Função gama)**.** *No exemplo* (1.6) *usamos a integral imprópria convergente*

$$\int_0^\infty e^{-x}\,\mathrm{d}x = -e^{-x}\Big|_0^\infty = 1$$

para concluir que

$$\int_0^\infty x^n e^{-x}\,\mathrm{d}x = n!$$

Com base nisso, a função gama é definida pela integral imprópria convergente

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}\,\mathrm{d}t$$

com o domínio ampliado para os escalares complexos.

Analogamente, a integral imprópria $\int_{-\infty}^b f(x)\,\mathrm{d}x$ é definida pelo limite

$$\lim_{a\to-\infty}\int_a^b f(x)\,\mathrm{d}x$$

Se ambas as integrais $\int_{-\infty}^c f(x)\,\mathrm{d}x$ e $\int_c^\infty f(x)\,\mathrm{d}x$ convergirem, então dizemos que $\int_{-\infty}^\infty f(x)\,\mathrm{d}x$ converge e

$$\int_{-\infty}^\infty f(x)\,\mathrm{d}x = \int_{-\infty}^c f(x)\,\mathrm{d}x + \int_c^\infty f(x)\,\mathrm{d}x$$

Mas se pelo menos uma dessas duas integrais divergir, então $\int_{-\infty}^\infty f(x)\,\mathrm{d}x$ diverge.

Teorema 2.18. *Seja $f(x)$ uma função integrável em um intervalo $[a,b]$, para qualquer $b \geq a$ e satisfaz a condição $\forall x \geq a \quad f(x) \geq 0$. Então $\int_a^\infty f(x)\,\mathrm{d}x$ converge se e somente se existir uma constante $M > 0$ tal que*

$$\int_a^b f(x)\,\mathrm{d}x \leq M \quad \forall b \geq a \tag{103}$$

Demonstração. Se (103) é satisfeita, então a integral parcial $\int_a^b f(x)\,\mathrm{d}x$ é uma função de b sempre crescente a partir de $x = a$ e é limitada superiormente. Denotando S como sendo o supremo de $\int_a^b f(x)\,\mathrm{d}x$, para qualquer $\epsilon > 0$ podemos encontrar um $\delta > a$ tal que $\left|S - \int_a^b f(x)\,\mathrm{d}x\right| < \epsilon \quad \forall b \geq \delta$, no que implica em sua convergência.

Se $\int_a^\infty f(x)\,\mathrm{d}x$ converge, então vamos fixar um M tal que $M > \int_a^\infty f(x)\,\mathrm{d}x$, e portanto $M > 0$. Uma vez que $\int_a^b f(x)\,\mathrm{d}x$ é sempre crescente em relação a b, segue que

$$\int_a^b f(x)\,\mathrm{d}x \leq \int_a^\infty f(x)\,\mathrm{d}x \quad \forall b \geq a$$

que implica (103). ■

Teorema 2.19. *Seja $f(x)$ uma função integrável em um intervalo $[a,b]$, para qualquer $b \geq a$ e $f(x)$ é não negativo $\forall x \geq a$. Se $g(x)$ é uma função tal que $0 \leq f(x) \leq g(x) \quad \forall x \geq a$ e a integral $\int_a^\infty g(x)\,\mathrm{d}x$ converge, então $\int_a^\infty f(x)\,\mathrm{d}x$ também converge e*

$$\int_a^\infty f(x)\,\mathrm{d}x \leq \int_a^\infty g(x)\,\mathrm{d}x$$

Demonstração. Uma vez que $0 \leq f(x) \leq g(x)$, então para qualquer partição temos $\int_a^b f(x)\,\mathrm{d}x \leq \int_a^b g(x)\,\mathrm{d}x$. Tomando o limite $b \to \infty$, obtemos

$$\int_a^\infty f(x)\,\mathrm{d}x \leq \int_a^\infty g(x)\,\mathrm{d}x$$

Por hipótese, $\int_a^\infty g(x)\,\mathrm{d}x$ converge e portanto, pelo teorema (2.18), $\int_a^\infty f(x)\,\mathrm{d}x$ também converge. ■

Teorema 2.20. *Sejam $f(x)$ e $g(x)$ duas funções tais que $\int_a^b f(x)\,\mathrm{d}x$ e $\int_a^b g(x)\,\mathrm{d}x$ existem para qualquer $b \geq a$ e satisfazem as condições*

$$\forall x \geq a \quad f(x) \geq 0 \quad g(x) > 0$$

Se $\lim_{x\to\infty} \frac{f(x)}{g(x)} = c$, onde $c \neq 0$, então ambas as integrais $\int_a^\infty f(x)\,\mathrm{d}x$ e $\int_a^\infty g(x)\,\mathrm{d}x$ convergem ou ambas divergem. E se $c = 0$, então $\int_a^\infty f(x)\,\mathrm{d}x$ só pode convergir se $\int_a^\infty g(x)\,\mathrm{d}x$ convergir.

Demonstração. Se $\lim_{x\to\infty} \frac{f(x)}{g(x)} = c$, e $c \neq 0$ então para algum $\delta \geq a$ as condições a seguir são satisfeitas:

$$0 < u \leq \frac{f(x)}{g(x)} \leq v \quad \forall x \geq \delta \tag{104}$$

onde $u \leq c$ e $v \geq c$. Sabendo que $g(x)$ é positivo, podemos multiplicar (104) por $g(x)$ sem inverter a ordem das desigualdades:

$$0 < ug(x) \leq f(x) \leq vg(x) \quad \forall x \geq \delta$$

Usando a propriedade da integral que conserva a ordem de desigualdade em funções não negativas e integrando os termos, chegamos na seguinte expressão:

$$0 < \int_d^b ug(x)\,\mathrm{d}x \leq \int_d^b f(x)\,\mathrm{d}x \leq \int_d^b vg(x)\,\mathrm{d}x \quad \forall x \geq \delta$$

onde $d \geq \delta$. No limite em que $b \to \infty$, as constantes u e v não interferem na convergência de $\int_d^\infty g(x)\,\mathrm{d}x$ e, portanto, se $\int_d^\infty vg(x)\,\mathrm{d}x$ convergir, então pelo teorema (2.19) $\int_d^\infty f(x)\,\mathrm{d}x$ também converge. Por raciocínio análogo, a divergência de uma das integrais também implica na divergência da outra integral, e o mesmo para convergência.

Se $c = 0$, então para algum $\delta \geq a$ as condições a seguir são satisfeitas:

$$0 \leq \frac{f(x)}{g(x)} \leq v \quad \forall x \geq \delta \tag{105}$$

Multiplicando por $g(x)$ e aplicando a integral:

$$0 \leq \int_d^b f(x)\,\mathrm{d}x \leq \int_d^b vg(x)\,\mathrm{d}x \quad \forall x \geq \delta \tag{106}$$

E disto apenas podemos concluir que se no limite $b \to \infty$ $\int_d^b g(x)\,\mathrm{d}x$ converge, então $\int_d^b f(x)\,\mathrm{d}x$ também converge. A sua contrapositiva também pode ser trivialmente deduzida.

E finalmente, a escolha da constante d não afeta a convergência ou divergência de qualquer uma das integrais trabalhadas acima, o que significa que os resultados também servem para $d = a$. ■

Definição 2.7 (Integral imprópria do tipo 2). *Seja $f(x)$ uma função integrável no intervalo $(a, b]$ e $I(x)$ a função*

$$I(x) = \int_x^b f(t)\,\mathrm{d}t$$

Dizemos que a integral imprópria do tipo 2 denotada por $\int_{a+}^b f(t)\,\mathrm{d}t$ converge se o limite

$$\lim_{x \to a+} I(x) = \lim_{x \to a+} \int_x^b f(t)\,\mathrm{d}t \tag{107}$$

existe e é finito. Se (107) existe e é igual a A, então denotamos

$$\int_{a+}^b f(t)\,\mathrm{d}t = A$$

A definição da integral $\int_a^{b-} f(t)\,\mathrm{d}t$ também é similar à integral anterior, onde $f(x)$ precisa ser integrável no intervalo $[a, b)$.

Se $f(x)$ é integrável em (a, b), $c \in (a, b)$ e ambas $\int_{a+}^c f(t)\,\mathrm{d}t$ e $\int_c^{b-} f(t)\,\mathrm{d}t$ convergem, então dizemos que $\int_{a+}^{b-} f(t)\,\mathrm{d}t$ também converge e

$$\int_{a+}^{b-} f(t)\,\mathrm{d}t = \int_{a+}^c f(t)\,\mathrm{d}t + \int_c^{b-} f(t)\,\mathrm{d}t$$

2.3 Convergência de séries e sequências de funções

Definição 2.8 (Convergência pontual). *Seja $\{f_n(x)\}$ uma sequência cujos termos possuem dependência à variável x. Diz-se que $\{f_n(x)\}$ converge pontualmente à função $f(x)$ se*

$$f(x) = \lim_{n \to \infty} f_n(x) \quad \textit{para todo } x \textit{ em que o limite converge}$$

Equivalentemente, $\{f_n(x)\}$ converge pontualmente para $f(x)$ em um intervalo I se para cada $x_0 \in I$ existir um N_{x_0} tal que

$$n \geq N_{x_0} \qquad \textit{implica} \qquad |f_n(x_0) - f(x_0)| < \epsilon$$

com $\epsilon > 0$ *arbitrário.*

A função $f(x)$ *se chama função limite da sequência* $\{f_n(x)\}$. *Também é dito que* $\{f_n(x)\}$ *converge pontualmente para* $f(x)$ *no domínio em que o limite converge.*

Exemplo 2.8 (Sequência de funções contínuas com função limite descontínua)**.** *Seja* $f_n(x) = x^n$ *cujo domínio é* $[0,1] \quad \forall n$. *No limite* $n \to \infty$, *a sequência* $\{f_n(x)\}$ *converge no intervalo* $[0,1]$ *para a função* $f(x)$*:*

$$f(x) = \begin{cases} 0 & \textit{se } 0 \leq x < 1 \\ 1 & \textit{se } x = 1 \end{cases}$$

Note que $f(x)$ *é descontínua em* $x = 1$ *apesar de todas as funções* $f_n(x)$ *serem contínuas no intervalo* $[0,1]$.

Exemplo 2.9 (Sequências tais que $\displaystyle\int_a^x \lim_{n\to\infty} f_n(t)\,\mathrm{d}t \neq \lim_{n\to\infty}\int_a^x f_n(t)\,\mathrm{d}t$)**.** *Dado* $f_n(t) = t^{-\frac{n+1}{n}}$, $x \neq 0$, $a = 1$, *temos*

$$\int_1^x \lim_{n\to\infty} f_n(t)\,\mathrm{d}t = \int_1^x \lim_{n\to\infty} t^{-\frac{n+1}{n}}\,\mathrm{d}t = \int_1^x \frac{1}{t}\,\mathrm{d}t = \ln|x|$$

que converge para todo $x > 0$. *Porém*

$$\lim_{n\to\infty}\int_1^x f_n(t)\,\mathrm{d}t = \lim_{n\to\infty}\int_1^x t^{-\frac{n+1}{n}}\,\mathrm{d}t = \lim_{n\to\infty}\frac{-n}{x^{\frac{1}{n}}}$$

que diverge para $-\infty$ *para qualquer* $x > 0$.

Similarmente, fazendo $f_n(x) = nx \cdot (1 - x^2)^n$ *e restringindo ao intervalo* $x \in [0.1]$*:*

$$\int_0^1 nx \cdot (1-x^2)^n\,\mathrm{d}x = -\frac{n}{2}\int_1^0 u^n\,\mathrm{d}u = \frac{n}{2}\cdot\left.\frac{u^{n+1}}{n+1}\right|_0^1 = \frac{1}{2}\cdot\frac{n}{n+1}$$

e portanto

$$\lim_{n\to\infty}\int_0^1 nx\cdot(1-x^2)^n\,\mathrm{d}x = \frac{1}{2}$$

Mas se aplicarmos o limite primeiro:

$$\lim_{n\to\infty} nx\cdot(1-x^2)^n = 0$$

sabendo que se $x \in [0,1]$, *então* $(1-x^2)^n$ *tende a zero mais rapidamente do que* n *a infinito.*

Definição 2.9 (Convergência uniforme)**.** *Uma sequência $\{f_n(x)\}$ diz-se convergir uniformemente para $f(x)$ em um intervalo I se para qualquer $\epsilon > 0$, existe um N tal que para todos os valores de $x \in I$*

$$n \geq N \quad \textit{implica} \quad |f_n(x) - f(x)| < \epsilon$$

Note que N satisfaz essas condições para todo x no intervalo I para um dado ϵ, enquanto que na convergência pontual cada x_0 possui seu próprio N_{x_0} associado ao ϵ.

Exemplo 2.10. *Colocando $f_n(x) = x^n$ na definição (2.9), temos que*

$$|x^n - f(x)| < \epsilon \Longleftrightarrow x^n - \epsilon < f(x) < x^n + \epsilon$$

onde

$$f(x) = \begin{cases} 0 & \textit{se } 0 \leq x < 1 \\ 1 & \textit{se } x = 1 \end{cases}$$

Agora vamos escolher $\epsilon = \dfrac{1}{2}$. E então para $x = \sqrt[n]{\dfrac{1}{2}}$, temos $x^n + \epsilon = 1$, $x^n - \epsilon = 0$ e portanto

$$0 < f(x) < 1$$

Mas não importa o quão grande seja n, o valor de $f\left(\sqrt[n]{\dfrac{1}{2}}\right)$ sempre será zero, o que viola essa desigualdade, e portanto $\{f_n(x)\}$ não pode convergir uniformemente.

Exemplo 2.11. *Seja $f_n(x) = nx(1-x^2)^n$. Sabemos que a sequência $\{f_n(x)\}$ converge pontualmente a zero no intervalo $[0,1]$. Portanto, a condição para que $\{f_n(x)\}$ convirja uniformemente em $[0,1]$ é de que dado $\epsilon > 0$ exista algum N tal que*

$$\left|nx(1-x^2)^n\right| < \epsilon \quad \forall n \geq N$$

para todo x em $[0,1]$.

A sua derivada é

$$\begin{aligned} f_n'(x) &= n(1-x^2)^n - 2n^2x^2(1-x^2)^{n-1} \\ &= n(1-x^2)^{n-1} \cdot [1 - x^2 - 2nx^2] \\ &= n(1-x^2)^{n-1} \cdot [1 - x^2(1+2n)] \end{aligned}$$

ou seja, $f_n'(x) = 0$ no intervalo $(0,1)$ quando

$$1 - x^2(1+2n) = 0$$

$$x = \sqrt{\frac{1}{2n+1}}$$

Substituindo em $f_n(x) = nx(1-x^2)^n$*, obtemos*

$$f_n(x_{max}) = \frac{n}{\sqrt{2n+1}}\left(1 - \frac{1}{2n+1}\right)^n$$

e então

$$\lim_{n\to\infty} f_n(x_{max}) = \left[\lim_{n\to\infty} \frac{n}{\sqrt{2n+1}}\right]\cdot\left[\lim_{n\to\infty}\left(1 - \frac{1}{2n+1}\right)^n\right] = +\infty\cdot\frac{1}{\sqrt{e}} = +\infty > \epsilon$$

ou seja, não importa o quão grande seja N*, existirá algum* $x \in (a,b)$ *tal que*

$$\left|nx(1-x^2)^n\right| > \epsilon \quad \forall n \geq N$$

o que significa que $\{f_n(x)\}$ *não converge uniformemente.*

Teorema 2.21 (Convergência uniforme e continuidade)**.** *Dada uma sequência* $\{f_n(x)\}$ *que converge uniformemente para* $f(x)$ *em um conjunto* S*, se cada função* $f_n(x)$ *é contínua em um ponto* $p \in S$*, então* $f(x)$ *também é contínua em* p*.*

Demonstração. Por hipótese, para qualquer $\epsilon > 0$ existe um inteiro N tal que $n \geq N$ implica

$$|f_n(x) - f(x)| < \frac{\epsilon}{3} \quad \forall x \in S \tag{108}$$

Dado que $f_n(x)$ é contínua em p, então há um intervalo de vizinhança $N(p)$ em S tal que

$$|f_N(x) - f_N(p)| < \frac{\epsilon}{3} \quad \forall x \in N(p) \cap S \tag{109}$$

Portanto $\forall x \in N(p) \cap S$ é válido

$$\begin{aligned}
|f(x) - f(p)| &= |f(x) - f_N(x) + f_N(x) - f_N(p) + f_N(p) - f(p)| \\
&\leq |f(x) - f_N(x)| + |f_N(x) - f_N(p)| + |f_N(p) - f(p)| \\
&< \underbrace{\frac{\epsilon}{3}}_{(108)} + \underbrace{\frac{\epsilon}{3}}_{(109)} + \underbrace{\frac{\epsilon}{3}}_{(109)} \\
&< \epsilon
\end{aligned}$$

o que implica que $f(x)$ é contínua em p. ■

Isso significa que se $f_n(x) = \sum_{k=1}^{n} u_k(x)$ e f_n converge pontualmente para f no intervalo $x \in I$, então dizemos que a seguinte série converge pontualmente e

$$f(x) = \sum_{k=1}^{\infty} u_k(x) \quad \forall x \in I$$

E se f_n converge uniformemente para f, então dizemos que a série converge uniformemente para f. Se cada termo u_k for contínua em p, então a sua soma

parcial também será contínua em p e portanto idem para f_n. Por consequência, pelo teorema (2.21), $f(x)$ também será contínua em p, no que implica no teorema a seguir.

Teorema 2.22. *Se $\sum u_k$ converge uniformemente para f em um conjunto S e cada termo u_k for contínuo em $p \in S$, então a função f também é contínua em p. Ou seja*

$$\lim_{x \to p} f(x) = \lim_{x \to p} \sum_{k=1}^{\infty} u_k(x) = \sum_{k=1}^{\infty} \lim_{x \to p} u_k(x)$$

Teorema 2.23. *Suponha que f_n converge uniformemente para f em um intervalo $[a, b]$. Dada a sequência $\{g_n\}$:*

$$g_n(x) = \int_a^x f_n(t)\,\mathrm{d}t$$

e a função

$$g(x) = \int_a^x f(t)\,\mathrm{d}t$$

então $g_n(x)$ converge uniformemente para $g(x)$ no intervalo $[a, b]$, ou seja

$$\lim_{n \to \infty} \int_a^x f_n(t)\,\mathrm{d}t = \int_a^x f(x)\,\mathrm{d}t = \int_a^x \lim_{n \to \infty} f_n(t)\,\mathrm{d}t$$

Demonstração. Como $f_n(t)$ converge uniformemente para $f(t)$, então

$$0 \leq |f_n(t) - f(t)| < \frac{\epsilon}{b-a} \quad \forall t \in [a, b] \quad \forall n \geq N \tag{110}$$

Integrando (110) em ambos os lados no intervalo $[a, b]$:

$$\int_a^b |f_n(t) - f(t)|\,\mathrm{d}t < \int_a^b \frac{\epsilon}{b-a}\,\mathrm{d}t = \epsilon$$

e então

$$\begin{aligned} |g_n(x) - g(x)| &= \left| \int_a^x f_n(t) - f(t)\,\mathrm{d}t \right| \\ &\leq \int_a^b |f_n(t) - f(t)|\,\mathrm{d}t \\ &< \epsilon \end{aligned}$$

o que prova que $g_n(x)$ converge uniformemente para $g(x)$. ■

Exemplo 2.12. *Dada a sequência $\{f_n(x)\}$, com $f_n(x) = x^{-\frac{n+1}{n}}$, ela não pode convergir uniformemente em nenhum intervalo $I \subseteq (0, \infty)$, uma vez que apesar de ser contínua em I para qualquer n, a integral do limite é diferente do limite da integral, o que significa que não obedece aos requisitos do teorema (2.23).*

Teorema 2.24. *Se* $f_n = \sum_{k=1}^{n} u_k$ *é uma série que converge uniformemente para* f *em um intervalo* $[a, b]$, *onde* u_k *é contínua em* $[a, b]$, *e para qualquer* $x \in [a, b]$ *são definidos*

$$g_n(x) = \sum_{k=1}^{n} \int_a^x u_k(t)\,\mathrm{d}t \quad g(x) = \int_a^x f(t)\,\mathrm{d}t$$

então $g_n(x)$ *converge uniformemente para* $g(x)$ *no intervalo* $[a, b]$, *ou seja*

$$\lim_{n\to\infty} \sum_{k=1}^{n} \int_a^x u_k(t)\,\mathrm{d}t = \int_a^x \lim_{n\to\infty} \sum_{k=1}^{n} u_k(t)\,\mathrm{d}t$$

Demonstração. Supondo u_k contínua em $[a, b]$, basta fazer $f_n(x) = \sum_{k=1}^{n} u_k(x)$ no teorema (2.23) e perceber que

$$g_n(x) = \int_a^x f_n(t)\,\mathrm{d}t = \int_a^x \sum_{k=1}^{n} u_k(x)\,\mathrm{d}t = \sum_{k=1}^{n} \int_a^x u_k(x)\,\mathrm{d}t$$

■

Teorema 2.25 (Teste M de Weierstrass)**.** *Se* $\sum u_n$ *é uma série que converge pontualmente para* f *em um conjunto* S *e existe uma série convergente* $\sum M_n$ *de constantes tal que*

$$0 \leq |u_n(x)| \leq M_n \quad \forall n \geq 1, \quad \forall x \in S$$

então a série $\sum u_n$ *converge uniformemente em* S.

Demonstração. Sabemos que pela convergência pontual

$$f(x) = \sum_{k=1}^{\infty} u_k(x)$$

e portanto

$$\left| f(x) - \sum_{k=1}^{n} u_k(x) \right| = \left| \sum_{k=n+1}^{\infty} u_k(x) \right| \leq \sum_{k=n+1}^{\infty} |u_k(x)|$$

Por hipótese, $|u_k(x)| \leq M_k$, no que implica

$$\left| f(x) - \sum_{k=1}^{n} u_k(x) \right| \leq \sum_{k=n+1}^{\infty} M_k$$

Dado que $\sum M_n$ converge, então é condição necessária

$$\lim_{k\to\infty} M_k = 0$$

E então podemos concluir que dado $\epsilon > 0$, existe algum N tal que $n \geq N$ implica

$$\sum_{k=n+1}^{\infty} M_k < \epsilon$$

e então

$$\left| f(x) - \sum_{k=1}^{n} u_k(x) \right| < \epsilon$$

o que significa que $\sum_{k=1}^{n} u_k(x)$ converge uniformemente para $f(x)$ no domínio $x \in S$. ■

2.3.1 Séries de potências

Uma série infinita do tipo

$$\sum_{n=0}^{\infty} a_n (z-a)^n \tag{111}$$

é chamada de série de potências centrada em a. Tanto z, a e os coeficientes a_n podem ser complexos. Vamos omitir a demonstração por enquanto mas é provado que existem somente três casos para a convergência dessa série:

- converge absolutamente em todo o $\mathbb{C}$;
- converge apenas no caso trivial $z = a$;
- converge absolutamente apenas dentro de uma região circular de raio r em torno de a.

Nesse último caso, não há qualquer garantia de convergência na borda em que $|z - a| = r$, podendo ser parcial, total ou até nenhum ponto.

Exemplo 2.13. *Fazendo* $a = 0$ *e* $a_n = \frac{1}{n!}$ *em* (111)*, obtemos a série de potências*

$$\sum_{n=0}^{\infty} \frac{z^n}{n!} \tag{112}$$

Para encontrarmos o seu raio de convergência r*, aplicaremos o teste da razão* (95) *na série absoluta e então impomos o critério de convergência em que o limite é menor do que 1:*

$$\lim_{n\to\infty} \left| \frac{b_{n+1}}{b_n} \right| = \lim_{n\to\infty} \left| \frac{z^{n+1}}{(n+1)!} \cdot \frac{n!}{z^n} \right| = \lim_{n\to\infty} \left| \frac{z}{n+1} \right| < 1$$

Uma vez que independentemente de z *o limite tenderá a zero, então o teste nos diz que a série* (112) *converge para todo o plano* $\mathbb{C}$*.*

Exemplo 2.14. *Fazendo* $a = 0$ *e* $a_n = n^2 3^n$ *em* (111)*, obtemos a seguinte série de potências:*

$$\sum_{n=0}^{\infty} n^2 3^n z^n = \sum_{n=0}^{\infty} n^2 (3z)^n \tag{113}$$

Desta vez aplicaremos o teste da raíz (93) *e estabelecer o limite menor do que* 1 *para encontrar o raio de convergência:*

$$\lim_{n\to\infty} \sqrt[n]{|b_n|} = \lim_{n\to\infty} \sqrt[n]{n^2(3|z|)^n} = \lim_{n\to\infty} \sqrt[n]{n}^2 \cdot 3|z| < 1$$

Quando $n \to \infty$*,* $\lim_{n\to\infty} \sqrt[n]{n}^2 = 1$ *e então a série converge absolutamente para qualquer* $z \in \mathbb{C}$ *tal que* $|z| < \dfrac{1}{3}$*. O raio de convergência é, portanto,* $r = \dfrac{1}{3}$*. A série diverge se* $|z| > \dfrac{1}{3}$*. No caso em que* $|z| = \dfrac{1}{3}$*, então a série absoluta de* (113) *se tornará*

$$\sum_{n=0}^{\infty} n^2$$

que diverge para qualquer ponto da fronteira, $|z| = \dfrac{1}{3}$*.*

Teorema 2.26. *Suponha que a série* $\sum a_n z^n$ *converge para algum* $z = z_1 \neq 0$*. Então*

1. *A série converge absolutamente para qualquer* z *tal que* $|z| < |z_1|$*.*
2. *A série converge uniformemente em qualquer círculo centrado em* 0 *e de raio* $r < |z_1|$*.*

Demonstração. Uma vez que $\sum_{n=0}^{\infty} a_n z_1^n$ converge, então é condição necessária

$$\lim_{n\to\infty} a_n z_1^n = 0$$

Isso significa que existe um N tal que $|a_n z_1^n| < 1 \quad \forall n \geq N$. Seja S um círculo de raio r, onde

$$0 < r < |z_1| \tag{114}$$

Se $z \in S$ e $n \geq N$, então $|z| \leq r$ e

$$|a_n z^n| = \left| a_n z_1^n \cdot \left(\frac{z}{z_1}\right)^n \right| = \underbrace{|a_n z_1^n|}_{<1} \cdot \left|\frac{z}{z_1}\right|^n < \frac{|z|^n}{|z_1|^n} \leq \frac{r^n}{|z_1|^n} = t^n$$

onde $t = \dfrac{r}{|z_1|}$. Dividindo (114) por $|z_1|$, temos que $0 < t < 1$. Isso significa que a série geométrica $\sum t^n$ converge. Fazendo $u_n(x) = a_n z^n$ e $M_n = t^n$ no teste M de Weierstrass (2.25) e sabendo que $\sum a_n z^n$ converge pontualmente uma vez que $|z| < |z_1|$, então $\sum a_n z^n$ converge uniformemente em S. Além disso, dado que $|a_n z^n| < t^n$, então $\sum a_n z^n$ também converge absolutamente para qualquer z tal que $|z| < |z_1|$. ■

Teorema 2.27 (Existência do círculo de convergência). *Suponha que a série $\sum a_n z^n$ converge para algum $z = z_1 \neq 0$ mas diverge para algum $z = z_2$. Então há um real positivo r tal que a série $\sum a_n z^n$ converge absolutamente se $|z| < r$ e diverge se $|z| > r$.*

Demonstração. Seja A o conjunto de todos os reais positivos $|z|$ tais que $\sum a_n z^n$ converge. Uma vez que $|z_1| \in A$, então A é um conjunto não vazio. Além disso, como $\sum a_n z_2^n$ diverge, pelo teorema (2.26) o real positivo $|z_2|$ não pode pertencer a A, cujo ínfimo é 0. Isso implica que $|z_2|$ é superior a qualquer elemento de A e portanto é um limite superior de A. É provado que se A é um conjunto não vazio de elementos reais e é limitado superiormente, então A tem um supremo. Dito isso, vamos denotar $\sup A = r$. Dado que $|z_1| > 0$, então $r \geq |z_1|$ implica $r > 0$. E então qualquer valor de z tal que $|z| > r$ a série $\sum a_n z^n$ diverge. E se $|z| < r$, então existe algum z_3 tal que $|z_3| = x$ com $|z| < x < r$. Nessas condições o teorema (2.26) garante que $\sum a_n z^n$ converge. ∎

A demonstração da existência do círculo de convergência para a série mais geral $\sum a_n (z-a)^n$ é feita repetindo as demonstrações anteriores com $z = z' - a$.

2.3.2 Funções representadas como séries reais de potências

Quando $z, a, a_n \in \mathbb{R}$ em $\sum a_n (z-a)^n$, o circulo de convergência de raio r se torna um intervalo de tamanho $2r$ centrado em a, denotado intervalo de convergência.

Dada uma série de potências, ela pode definir uma função real f da seguinte forma:

$$f(x) = \sum_{n=0}^{\infty} a_n (x-a)^n$$

Por $f(x)$ poder ser representada como uma série de potências, essa função apresenta algumas propriedades interessantes.

Teorema 2.28. *Seja f a função definida (ou representada) pela expressão*

$$f(x) = \sum_{n=0}^{\infty} a_n (x-a)^n \tag{115}$$

cujo domínio é o intervalo $(a-r, a+r)$. Então f é contínua nesse intervalo e a integral de f em qualquer subintervalo de $(a-r, a+r)$ pode ser calculada pela série infinita da integral de cada termo $a_n(x-a)^n$:

$$\int_b^c f(x)\,\mathrm{d}x = \sum_{n=0}^{\infty} \left[\int_b^c a_n (x-a)^n \,\mathrm{d}x \right] = \sum_{n=0}^{\infty} \frac{a_n}{n+1} \cdot (x-a)^{n+1}$$

onde $b, c \in (a-r, a+r)$.

Demonstração. O teorema (2.26) garante que $\sum a_n (x-a)^n$ converge absolutamente no intervalo $(a-r, a+r)$ e que se R é um real positivo tal que $0 < R < r$, então $\sum a_n (x-a)^n$ converge uniformemente no intervalo $[a-R, a+R]$. Sabendo

que os termos $a_n(x-a)^n$ são contínuos em todo o $\mathbb{R}$, mais especificamente em $(a-r, a+r)$, podemos usar o teorema (2.22). Além disso, se $[b, c]$ é um subintervalo de $[a-R, a+R]$, então podemos usar o teorema (2.24) com os limites de integração em b e c. ■

Teorema 2.29. *Dada a função* (115) *em um intervalo de convergência* $(a-r, a+r)$*, ela obedece as seguintes propriedades:*

1. *O raio de convergência da série obtida derivando* (115) *termo a termo,* $\sum_{n=1}^{\infty} na_n(x-a)^{n-1}$*, também é* r*.*
2. *A derivada* $f'(x)$ *é definida para todo o* x *no intervalo* $(a-r, a+r)$ *e é dado por*

$$f'(x) = \sum_{n=1}^{\infty} na_n(x-a)^{n-1} \tag{116}$$

Demonstração. Suponhamos que $a = 0$. Então o intervalo de interesse é $(-r, r)$. Escolhemos um x tal que $0 < x < r$. Existe um positivo h pequeno o suficiente para que satisfaça $0 < x < x+h < r$. Então pelo teorema (2.26), ambos $f(x)$ e $f(x+h)$ convergem absolutamente. Por conseguinte, podemos fazer

$$\frac{f(x+h)-f(x)}{h} = \sum_{n=0}^{\infty} \frac{a_n}{h}((x+h)^n - x^n) \tag{117}$$

Pelo teorema (2.2) da combinação linear, (117) também converge absolutamente. Pelo teorema do valor médio, existe um valor c_n satisfazendo $x < c_n < x+h$ tal que

$$\frac{(x+h)^n - x^n}{h} = \frac{d}{dx}[x^n]\bigg|_{c_n} = nc_n^{n-1}$$

e portanto

$$\sum_{n=0}^{\infty} \frac{a_n}{h}((x+h)^n - x^n) = \sum_{n=0}^{\infty} na_n c_n^{n-1} = \sum_{n=1}^{\infty} na_n c_n^{n-1} \tag{118}$$

Uma vez que (117) é absolutamente convergente, então a série (118) também converge absolutamente. Sabendo que

$$0 < x < c_n \implies \left|na_n x^{n-1}\right| < \left|na_n c_n^{n-1}\right|$$

então a série $\sum \left|na_n x^{n-1}\right|$ é limitada superiormente:

$$\sum_{n=1}^{\infty} \left|na_n x^{n-1}\right| < \sum_{n=1}^{\infty} \left|na_n c_n^{n-1}\right|$$

o que implica que $\sum na_n x^{n-1}$ tem um raio de convergência de pelo menos r.

Se fixarmos algum $x \neq 0$, então existe algum N tal que para todo $n \geq N$ temos

$$1 < \left|\frac{n}{x}\right|$$

e portanto, para todo $x \neq 0$, existe um N tal que

$$|a_n x^n| < |a_n x^n| \cdot \left|\frac{n}{x}\right| = \left|n a_n \cdot x^{n-1}\right|$$

e consequentemente

$$\sum_{n=N}^{\infty} |a_n x^n| < \sum_{n=N}^{\infty} \left|n a_n \cdot x^{n-1}\right|$$

Uma vez que para qualquer $x > r$ a série $\sum |a_n x^n|$ diverge, então $\sum \left|n a_n x^{n-1}\right|$ também diverge, o que significa que o raio de convergência de $\sum \left|n a_n x^{n-1}\right|$ não pode ser maior do que r.

Agora seja $g(x)$ a função

$$g(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}$$

Devemos provar que $g(x) = f'(x)$. Aplicando o teorema (2.28) em $g(x)$, podemos integrar termo a termo:

$$\int_0^x g(t)\,\mathrm{d}t = \sum_{n=1}^{\infty} \int_0^x n a_n t^{n-1}\,\mathrm{d}t = \sum_{n=1}^{\infty} a_n t^n \bigg|_0^x = \sum_{n=1}^{\infty} a_n x^n$$

$$\sum_{n=1}^{\infty} a_n x^n = \left[\sum_{n=0}^{\infty} a_n x^n\right] - a_0 = f(x) - a_0$$

Isso significa que $f(x)$ é primitiva de $g(x)$, no que $f'(x)$ existe e $f'(x) = g(x)$.

A demonstração para $a \neq 0$ é análoga, substituindo $x - a$ por x' e notando que esse caso é a série com $a = 0$ transladada em a para o sentido positivo do eixo x. Consequentemente o raio de convergência é invariante. ■

Se F é uma função definida pela expressão

$$F(x) = C + \sum_{n=0}^{\infty} \frac{a_n (x-a)^{n+1}}{n+1} \tag{119}$$

então pelo teorema (2.29) sua derivada é dada por

$$F'(x) = \sum_{n=0}^{\infty} a_n (x-a)^n$$

com o raio de convergência igual a r. Ou seja, as funções (115), (116) e (119) possuem o mesmo raio de convergência. A derivação ou integração das séries de potência preservam, portanto, o raio de convergência e podemos derivar ou integrar quantas vezes for o desejado.

Exemplo 2.15 (Arco tangente)**.** *Substituindo* x *por* $-x^2$ *em* (89)*, obtemos*

$$\sum_{n=0}^{\infty}(-x^2)^n = \sum_{n=0}^{\infty}(-1)^n x^{2n} = \frac{1}{1+x^2} \qquad se \qquad \left|-x^2\right| < 1 \implies |x| < 1$$

Podemos usar o teorema (2.28) *e integrar ambos os lados:*

$$\sum_{n=0}^{\infty}\int_0^{x'}(-1)^n x^{2n}\,\mathrm{d}x = \int_0^{x'}\frac{1}{1+x^2}\,\mathrm{d}x \qquad se \qquad |x| < 1$$

$$\sum_{n=0}^{\infty}\frac{(-1)^n}{2n+1}x^{2n+1} = \arctan(x)$$

desde que $|x| < 1$*.*

Exemplo 2.16 (Logaritmo e seu intervalo de convergência como série infinita)**.** *Substituindo* x *por* $1-x$ *em* (89)*, obtemos*

$$\sum_{n=0}^{\infty}(1-x)^n = \sum_{n=0}^{\infty}(-(x-1))^n = \sum_{n=0}^{\infty}(-1)^n(x-1)^n = \frac{1}{x}$$

onde o intervalo de convergência é dado por $|1-x| < 1$*, ou seja,* $0 < x < 2$*. Integrando ambos os lados resulta*

$$\sum_{n=0}^{\infty}\int_1^{x'}(-1)^n(x-1)^n\,\mathrm{d}x = \int_1^{x'}\frac{1}{x}\,\mathrm{d}x$$

$$\sum_{n=0}^{\infty}\frac{(-1)^n}{n+1}(x-1)^{n+1} = \ln(x)$$

desde que $0 < x < 2$*.*

Exemplo 2.17. *Seja* f *a função*

$$f(x) = \sum_{n=0}^{\infty}a_n(x-a)^n$$

Derivando k *vezes a função* f *obtemos*

$$f^{(k)}(x) = \sum_{n=k}^{\infty}\frac{n!}{(n-k)!}a_n(x-a)^{n-k}$$

Se fixarmos $x = a$*, então todos os termos da série se anularão exceto o de ordem zero:*

$$f^{(k)}(a) = \frac{k!}{0!}\cdot a_k$$

e concluímos que se $k \geq 1$, então

$$a_k = \frac{f^{(k)}(a)}{k!}$$

Essa fórmula também é válida para $k = 0$ se definirmos $f^{(0)}(a) = f(a)$. Logo, podemos afirmar que

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n \tag{120}$$

Uma vez que o valor de $f^{(n)}(a)$ só depende do comportamento de f na vizinhança de a, podemos enunciar o seguinte teorema:

Teorema 2.30 (Singularidade das séries de potências). *Se $\sum a_n(x-a)^n$ e $\sum b_n(x-a)^n$ possuem a mesma soma para qualquer ponto na vizinhança de a, então ambos a_n e b_n possuem o mesmo valor $\dfrac{f^{(n)}(a)}{n!}$ para todo $n \geq 0$.*

Dada a variedade de funções que podem ter representação em série de potências em um certo intervalo, suas propriedades e a frequência com que aparecem, é útil dar um nome a tais funções.

Definição 2.10 (Funções reais analíticas). *Uma função real f é dita analítica em um intervalo I se $\forall x \in I$ existe uma série de potências que converge para $f(x)$, ou seja*

$$f(x) = \sum_{n=0}^{\infty} a_n(x-a)^n \quad \forall x \in I$$

2.3.3 Série de Taylor de funções

A fórmula (120) indica que se f é analítica, então a sua soma parcial é justamente o polinômio de Taylor de $f(x)$ centrada em $x = a$, que converge pontualmente para f no intervalo $(a-r, a+r)$ e converge uniformemente em qualquer subintervalo de $(a-r, a+r)$.

Por outro lado, dada uma função f qualquer, existe a possibilidade de f ser analítica em um certo intervalo, como visto nos exemplos (2.15) e (2.16).

Suponha que f seja uma função cujas derivadas de qualquer ordem existem em um intervalo aberto nas vizinhanças de a. A função f é dita infinitamente diferenciável nesse intervalo. Então podemos construir a série de potências

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n \tag{121}$$

A expressão (121) é chamada de série de Taylor gerada por f em $x = a$. Entretanto isso não significa que necessariamente (121) converge uniformemente para f em todo o seu domínio. Aliás, existe a possibilidade dessa série nem sequer convergir se $x \neq a$. O seguinte teorema nos dá uma condição suficiente para a série de Taylor convergir pontualmente para f. A sua demonstração será omitida mas pode ser vista na seção 11.10 de [1].

Teorema 2.31. *Seja f uma função infinitamente diferenciável em um intervalo aberto $I = (a - r, a + r)$. Suponha que existe uma constante positiva A tal que*

$$\left| f^{(n)}(x) \right| \leq A^n \quad \forall n \geq 1$$

seja válido para todo $x \in I$. Então a série de Taylor gerada por f em $x = a$ converge pontualmente para $f(x)$ em todo $x \in I$.

2.4 Introdução às Equações Diferenciais Parciais

Dada uma função f de duas ou mais variáveis, uma equação envolvendo f e suas derivadas parciais é chamada de equação diferencial parcial (EDP).

Uma EDP é dita homogênea se o termo da equação que não depende de f nem de suas derivadas parciais é zero. Caso contrário, ela é dita não-homogênea. Além disso, dizemos que a ordem de uma EDP é n quando o maior de todas as ordens das derivadas é n.

Exemplo 2.18. *A equação diferencial parcial mais simples é a dada pela equação de primeira ordem*

$$\frac{\partial f(x,y)}{\partial x} = 0 \tag{122}$$

que é também homogênea.

Exemplo 2.19 (Equação de Laplace bidimensional)**.** *A equação de Laplace para duas dimensões é de segunda ordem e é dada pela equação*

$$\frac{\partial^2 f(x,y)}{\partial x^2} + \frac{\partial^2 f(x,y)}{\partial y^2} = 0 \tag{123}$$

Sabendo que a derivada ordinária é um operador linear, então a derivada parcial também é um operador linear e, dessa forma, ambas as EDPs anteriores são homogêneas lineares, ou seja, podem ser representadas por um operador linear $L(f) = 0$, onde $L = \dfrac{\partial}{\partial x}$ e $L = \dfrac{\partial^2}{\partial x^2} + \dfrac{\partial^2}{\partial y^2}$ respectivamente.

A representação $L(f) = 0$ também nos diz que as soluções dessa EDP formam um espaço vetorial devido ao princípio da superposição, uma vez que se f e g são soluções da equação e $a, b \in \mathbb{C}$, então

$$L(af + bg) = aL(f) + bL(g) = 0$$

ou seja, a combinação linear $af + bg$ das soluções também é uma solução da equação diferencial.

A equação (122) nos diz que $f(x, y)$ não depende de x e portanto uma solução dessa equação diferencial parcial é uma função arbitrária de y:

$$f(x, y) = g(y)$$

Dada a arbitrariedade de g, podemos escolher infinitas soluções linearmente independentes e, portanto, o espaço das soluções tem dimensão infinita.

2.4.1 EDPs lineares de primeira ordem

Uma equação diferencial parcial linear de primeira ordem de duas variáveis pode ser expressa na forma

$$a(x,y)\cdot\frac{\partial f(x,y)}{\partial x}+b(x,y)\cdot\frac{\partial f(x,y)}{\partial y}+c(x,y)\cdot f(x,y)=g(x,y) \tag{124}$$

onde as funções $a(x,y)$, $b(x,y)$, $c(x,y)$ e $g(x,y)$ são denominadas coeficientes da EDP. No exemplo seguinte, trataremos uma versão mais simplificada de (124) onde os coeficientes são constantes para todo x e y.

Exemplo 2.20. *Seja a EDP*

$$3\cdot\frac{\partial f(x,y)}{\partial x}+2\cdot\frac{\partial f(x,y)}{\partial y}=0 \tag{125}$$

Essa expressão pode ser reescrita como um produto escalar:

$$(3\hat{\mathbf{i}}+2\hat{\mathbf{j}})\cdot\nabla f(x,y)=0$$

Isso significa que $\nabla f(x,y)$ é ortogonal ao vetor $(3\hat{\mathbf{i}}+2\hat{\mathbf{j}})$.

Agora sejam x e y quaisquer valores que satisfaçam

$$f(x,y)=C$$

para algum C pertencente à imagem de f. Isso significa que para esses valores

$$\nabla f(x,y)=\left(\frac{\partial f(x,y)}{\partial x},\frac{\partial f(x,y)}{\partial y}\right)=(0,0)$$

ou seja

$$(x\hat{\mathbf{i}}+y\hat{\mathbf{j}})\cdot\nabla f(x,y)=(x\hat{\mathbf{i}}+y\hat{\mathbf{j}})\cdot 0=0$$

Isso significa que $\nabla f(x,y)$ também é ortogonal às curvas de nível de f. Uma vez que esses produtos escalares ocorrem no $\mathbb{R}^2$, consequentemente as curvas de nível $f(x,y)=C$ são paralelas aos múltiplos de $(3\hat{\mathbf{i}}+2\hat{\mathbf{j}})$. Os múltiplos de um vetor constituem uma reta e portanto todas as curvas de nível $f(x,y)=C$ são retas paralelas ao vetor $(3\hat{\mathbf{i}}+2\hat{\mathbf{j}})$. Ou seja, as curvas de nível de f são a família de retas

$$f(x,y)=2x-3y=C$$

Isso significa que se a escolha de x e y for tal que $2x-3y$ seja constante, então $f(x,y)$ também o é. Isso nos diz que

$$f(x,y)=g(2x-3y) \tag{126}$$

para alguma função g. As suas derivadas parciais são

$$\frac{\partial}{\partial x}g(2x-3y)=2\cdot g'(2x-3y)\quad \frac{\partial}{\partial y}g(2x-3y)=-3\cdot g'(2x-3y)$$

e portanto

$$3 \cdot \frac{\partial}{\partial x} g(2x-3y) + 2 \cdot \frac{\partial}{\partial y} g(2x-3y) = 3 \cdot 2 \cdot g'(2x-3y) + 2 \cdot (-3) \cdot g'(2x-3y)$$
$$= 0$$

o que significa que qualquer função derivável $g(2x-3y)$ *é solução de* (125).

Entretanto existe a possibilidade de haver outras soluções que não sejam $g(2x-3y)$. *Para mostrar que toda solução de* (125) *resulta em* (126), *fazemos a mudança de variáveis*

$$x = Au + Bv \quad y = Cu + Dv \tag{127}$$

para poder reescrever f *como uma função* h *de* u *e* v*:*

$$f(x,y) = f(Au+Bv, Cu+Dv)$$

Calculando sua derivada parcial em u*:*

$$\frac{\partial}{\partial u} f(x(u,v), y(u,v)) = \frac{\partial f}{\partial x} \cdot \frac{\partial x}{\partial u} + \frac{\partial f}{\partial y} \cdot \frac{\partial y}{\partial u} = \frac{\partial f}{\partial x} \cdot A + \frac{\partial f}{\partial y} \cdot C$$

Se escolhermos $A = 3$ *e* $C = 2$, *então podemos usar a equação* (125) *e concluir que* $\dfrac{\partial f}{\partial u} = 0$ *e*

$$f(x(u,v), y(u,v)) = g(v)$$

para alguma função g. *Segue de* (127) *que*

$$2x - 3y = 6u - 6u + (2B-3D)v$$
$$= (2B-3D)v$$

Agora escolhemos B *e* D *tais que* $2B - 3D = 1$ *para obter* $v = 2x - 3y$ *e portanto*

$$f(x(u,v), y(u,v)) = g(v) \implies f(x,y) = g(2x-3y)$$

como queríamos demonstrar.

O exemplo anterior pode servir de demonstração para o seguinte teorema:

Teorema 2.32. *Se* g *é uma função diferenciável em* $\mathbb{R}$ *e* f *uma função de duas variáveis diferenciável em* $\mathbb{R}^2$ *definida pela equação*

$$f(x,y) = g(bx - ay) \tag{128}$$

onde pelo menos a *ou* b *são diferentes de zero, então* f *é solução da EDP*

$$a \cdot \frac{\partial f(x,y)}{\partial x} + b \cdot \frac{\partial f(x,y)}{\partial y} = 0 \tag{129}$$

em todo o $\mathbb{R}^2$. *Além disso, toda solução de* (129) *pode ser escrito na forma* (128).

Exemplo 2.21. *Seja f uma função de derivada contínua que pode ser expressa na forma*

$$f(x,y) = u(xy)$$

Então

$$\frac{\partial}{\partial x} f(x,y) = u'(xy) \cdot \frac{\partial (xy)}{\partial x} = y \cdot u'(xy)$$
$$\frac{\partial}{\partial y} f(x,y) = u'(xy) \cdot \frac{\partial (xy)}{\partial y} = x \cdot u'(xy)$$

e portanto $f(x,y)$ satisfaz a EDP

$$x \frac{\partial}{\partial x} f(x,y) - y \frac{\partial}{\partial y} f(x,y) = 0$$

2.4.2 EDPs lineares de segunda ordem

Uma equação diferencial parcial linear de segunda ordem possui a forma (124) acrescida dos termos

$$d(x,y)\frac{\partial^2}{\partial x^2} f(x,y) + j(x,y)\frac{\partial^2}{\partial y^2} f(x,y) + h(x,y)\frac{\partial^2}{\partial x \partial y} f(x,y)$$

Exemplo 2.22. *Seja a EDP*

$$\frac{\partial^2}{\partial x \partial y} f(x,y) = 0 \tag{130}$$

onde f é uma função duas vezes diferenciável. Pelo teorema de Clairaut-Schwarz

$$\frac{\partial}{\partial x}\left(\frac{\partial}{\partial y} f(x,y)\right) = \frac{\partial}{\partial y}\left(\frac{\partial}{\partial x} f(x,y)\right)$$

Então a equação (130) *nos diz que*

$$\frac{\partial}{\partial y} f(x,y) = g(y) \quad e \quad \frac{\partial}{\partial x} f(x,y) = h(x)$$

Integrando a primeira equação em relação a y e a segunda a x, obtemos

$$f(x,y) = G(y) + C_1(x) \quad e \quad f(x,y) = H(x) + C_2(y)$$

Igualando as expressões, concluímos que

$$G(y) + C_1(x) = H(x) + C_2(y) \implies C_1(x) = H(x) \quad e \quad C_2(y) = G(y)$$

e portanto as soluções de (130) *podem ser expressas como*

$$f(x,y) = H(x) + G(y)$$

Exemplo 2.23 (Equação unidimensional das cordas)**.** *A representação de uma onda unidimensional por uma função $f(x,t)$ duas vezes diferenciável obedece a EDP*

$$\frac{1}{v^2}\frac{\partial^2}{\partial t^2}f(x,t) = \frac{\partial^2}{\partial x^2}f(x,t) \tag{131}$$

onde v é a velocidade de propagação da onda. As condições iniciais deste problema são de que a corda se encontre em um estado inicial $F(x)$ em $t=0$, ou seja

$$f(x,0) = F(x)$$

e que a velocidade transversal inicial em cada ponto x é dada pela equação

$$\left.\frac{\partial}{\partial t}f(x,t)\right|_{t=0} = G(x)$$

Fazemos então a mudança de variáveis

$$x = Au + Bw \quad t = Cu + Dw$$

Escolhemos $A = \frac{1}{2}$, $B = \frac{1}{2}$, $C = \frac{1}{2v}$ e $D = -\frac{1}{2v}$;

$$x = \frac{u}{2} + \frac{w}{2} \quad t = \frac{u}{2v} - \frac{w}{2v} \tag{132}$$

$$g(u,w) = f\left(\frac{u}{2} + \frac{w}{2}, \frac{u}{2v} - \frac{w}{2v}\right)$$

Vamos mostrar que essa escolha satisfaz

$$\frac{\partial^2}{\partial u \partial w}g(u,w) = 0 \tag{133}$$

Calculando suas derivadas parciais em u e w:

$$\frac{\partial}{\partial u}g = \frac{\partial f}{\partial x}\cdot\frac{\partial x}{\partial u} + \frac{\partial f}{\partial t}\cdot\frac{\partial t}{\partial u} = \frac{1}{2}\frac{\partial f}{\partial x} + \frac{1}{2v}\frac{\partial f}{\partial t}$$

$$\frac{\partial^2}{\partial u \partial w}g(u,w) = \frac{\partial}{\partial x}\left[\frac{1}{2}\frac{\partial f}{\partial x} + \frac{1}{2v}\frac{\partial f}{\partial t}\right]\cdot\frac{\partial x}{\partial w} + \frac{\partial}{\partial t}\left[\frac{1}{2}\frac{\partial f}{\partial x} + \frac{1}{2v}\frac{\partial f}{\partial t}\right]\cdot\frac{\partial t}{\partial w}$$

$$\frac{\partial^2}{\partial u \partial w}g(u,w) = \left[\frac{1}{2}\frac{\partial^2 f}{\partial x^2} + \frac{1}{2v}\frac{\partial^2 f}{\partial x \partial t}\right]\cdot\frac{1}{2} + \left[\frac{1}{2}\frac{\partial^2 f}{\partial t \partial x} + \frac{1}{2v}\frac{\partial^2 f}{\partial t^2}\right]\cdot -\frac{1}{2v}$$

$$\frac{\partial^2}{\partial u \partial w}g(u,w) = \frac{1}{4}\frac{\partial^2 f}{\partial x^2} + \frac{1}{4v}\frac{\partial^2 f}{\partial x \partial t} - \frac{1}{4v}\frac{\partial^2 f}{\partial t \partial x} - \frac{1}{4v^2}\frac{\partial^2 f}{\partial t^2}$$

$$\frac{\partial^2}{\partial u \partial w}g(u,w) = \frac{1}{4}\cdot\left[\frac{\partial^2}{\partial x^2}f(x,t) - \frac{1}{v^2}\frac{\partial^2}{\partial t^2}f(x,t)\right]$$

Usando a equação (131)*, fica demonstrado* (133)*.*

Vimos no exemplo (2.22) que as soluções de (133) podem ser escritas como

$$g(u, w) = \varphi_1(u) + \varphi_2(w)$$

Isolando u e v nas equações (132), temos que

$$u = x + vt \quad e \quad w = x - vt$$

e portanto

$$f(x,t) = \varphi_1(x + vt) + \varphi_2(x - vt) \tag{134}$$

Fazendo $t = 0$, temos que

$$F(x) = \varphi_1(x) + \varphi_2(x) \tag{135}$$

Derivando (134) em relação a t obtemos

$$\frac{\partial}{\partial t} f(x,t) = v\varphi_1'(x + vt) - v\varphi_2'(x - vt)$$

e em $t = 0$:

$$G(x) = v\varphi_1'(x) - v\varphi_2'(x)$$

$$\frac{1}{v}G(x) = \varphi_1'(x) - \varphi_2'(x) \tag{136}$$

Derivando (135) em relação a x, temos que

$$F'(x) = \varphi_1'(x) + \varphi_2'(x) \tag{137}$$

Fazendo combinações lineares apropriadas de (136) e (137), chegamos ás expressões

$$\varphi_1'(x) = \frac{1}{2}\left[F'(x) + \frac{1}{v}G(x)\right]$$

$$\varphi_2'(x) = \frac{1}{2}\left[F'(x) - \frac{1}{v}G(x)\right]$$

Integrando ambas as equações em relação a x no intervalo 0 a x', concluímos que

$$\varphi_1(x) - \varphi_1(0) = \frac{F(x) - F(0)}{2} + \frac{1}{2v}\int_0^x G(s)\,\mathrm{d}s \tag{138}$$

$$\varphi_2(x) - \varphi_2(0) = \frac{F(x) - F(0)}{2} - \frac{1}{2v}\int_0^x G(s)\,\mathrm{d}s \tag{139}$$

Substituindo x por $x + vt$ em (138) e por $x - vt$ em (139) e somando ambas as equações, encontramos

$$f(x,t) - (\varphi_1(0) + \varphi_2(0)) = \frac{F(x+vt) - F(0)}{2} + \frac{F(x-vt) - F(0)}{2} + \frac{1}{2v}\int_0^{x+vt} G(s)\,\mathrm{d}s - \frac{1}{2v}\int_0^{x-vt} G(s)\,\mathrm{d}s$$

Sabemos que $F(0) = \varphi_1(0) + \varphi_2(0)$ *e portanto*

$$f(x,t) = \frac{F(x+vt) + F(x-vt)}{2} + \frac{1}{2v}\int_{x-vt}^{x+vt} G(s)\,\mathrm{d}s \tag{140}$$

Note que (140) *também é solução de* (131). *Essa equação é a solução de D'Alembert da EDP da onda unidimensional.*

Exemplo 2.24 (Laplaciano)**.** *Seja* $f(x,y,z)$ *uma função duas vezes diferenciável cujo valor só depende da distância do ponto* (x,y,z) *à origem, ou seja*

$$f(x,y,z) = g(r) \quad \text{onde} \quad r = \sqrt{x^2+y^2+z^2}$$

Segue que para $(x,y,z) \neq (0,0,0)$

$$\frac{\partial f}{\partial x} = \frac{dg(r)}{dr}\cdot\frac{\partial r}{\partial x}$$

$$\frac{\partial r}{\partial x} = \frac{x}{\sqrt{x^2+y^2+z^2}}$$

$$\frac{\partial f}{\partial x} = g'(r)\cdot\frac{x}{\sqrt{x^2+y^2+z^2}}$$

$$\frac{\partial^2 f}{\partial x^2} = \frac{dg'(r)}{dr}\cdot\frac{\partial r}{\partial x}\cdot\frac{x}{\sqrt{x^2+y^2+z^2}} + g'(r)\cdot\frac{\partial}{\partial x}\frac{x}{\sqrt{x^2+y^2+z^2}}$$

$$\frac{\partial^2 f}{\partial x^2} = g''(r)\cdot\frac{x^2}{x^2+y^2+z^2} + g'(r)\cdot\frac{\sqrt{x^2+y^2+z^2} - x\cdot\frac{x}{\sqrt{x^2+y^2+z^2}}}{x^2+y^2+z^2}$$

$$\frac{\sqrt{x^2+y^2+z^2} - x\cdot\frac{x}{\sqrt{x^2+y^2+z^2}}}{x^2+y^2+z^2} = \frac{y^2+z^2}{(x^2+y^2+z^2)^{3/2}}$$

$$\frac{\partial^2 f}{\partial x^2} = g''(r)\cdot\frac{x^2}{x^2+y^2+z^2} + g'(r)\cdot\frac{y^2+z^2}{(x^2+y^2+z^2)^{3/2}}$$

Analogamente,

$$\frac{\partial^2 f}{\partial y^2} = g''(r)\cdot\frac{y^2}{x^2+y^2+z^2} + g'(r)\cdot\frac{x^2+z^2}{(x^2+y^2+z^2)^{3/2}}$$

$$\frac{\partial^2 f}{\partial z^2} = g''(r)\cdot\frac{z^2}{x^2+y^2+z^2} + g'(r)\cdot\frac{x^2+y^2}{(x^2+y^2+z^2)^{3/2}}$$

e portanto

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = g''(r)\cdot\frac{x^2+y^2+z^2}{x^2+y^2+z^2} + g'(r)\cdot 2\frac{x^2+y^2+z^2}{(x^2+y^2+z^2)^{3/2}}$$

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = g''(r) + \frac{2}{r}g'(r) \tag{141}$$

Agora seja a EDP laplaciana

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = 0 \tag{142}$$

De (141), *segue que*

$$g''(r) + \frac{2}{r}g'(r) = 0 \tag{143}$$

para todo $r > 0$. *Uma de suas soluções é a função constante* $g_0(r) = C$. *Para encontrar a outra solução linearmente independente, fazemos* $h(r) = g'(r)$ *e então a equação diferencial ordinária (EDO)* (143) *se torna*

$$h'(r) + \frac{2}{r}h(r) = 0$$

E então

$$\frac{h'(r)}{h(r)} = -\frac{2}{r}$$

$$\frac{d}{dr}\ln(h(r)) = -\frac{2}{r}$$

Integrando de ambos os lados:

$$\ln(h(r)) = -2\ln(r) + A_0$$

$$= \ln\frac{A_1}{r^2}$$

$$h(r) = \frac{A_1}{r^2}$$

Por fim, integrando a última equação podemos encontrar $g(r)$*:*

$$g(r) = -\frac{A_1}{2r} + C$$

Ou seja, a solução da EDP (142) *é a função* f *definida por*

$$f(x, y, z) = -\frac{A}{2} \cdot \frac{1}{\sqrt{x^2 + y^2 + z^2}} + C$$

onde A *e* C *são constantes arbitrárias.*

2.4.3 Problemas de valor inicial e problemas de valor de contorno

Dada a arbitrariedade das soluções de uma EDP geral, podemos escolher uma solução que satisfaça alguma condição específica. A quantidade de condições independentes entre si que podemos impor para uma solução depende da dimensão do espaço solução dessa EDP. E o tipo dessa condição pode ser um problema de valor inicial ou de contorno.

Exemplo 2.25 (Problema de valor inicial)**.** *Seja a EDP do exemplo* (130)*, cuja solução é dada por*

$$f(x, y) = H(x) + G(y)$$

Como o espaço solução é de dimensão infinita, podemos considerar conveniente uma solução que seja múltipla de qualquer outra função específica, como

$$f(x, y) = C \cdot \sin(ax) + D \cdot \cos(by)$$

E como a EDP é de segunda ordem, podemos escolher um valor arbitrário em um ponto qualquer do $\mathbb{R}^2$ *assim como sua derivada parcial naquele ponto, como* $f(0,0) = 0$ *e* $\dfrac{\partial f}{\partial x}(0,0) = 1$*, o que resultaria em*

$$f(0,0) = D = 0 \implies f(x, y) = C \cdot \sin(ax)$$

$$\frac{\partial f}{\partial x} = C \cdot a \cdot \cos(ax)$$

$$\frac{\partial f}{\partial x}(0,0) = 1 \implies 1 = C \cdot a \implies C = \frac{1}{a}$$

ou seja, uma solução desse problema de valor inicial é dado por

$$f(x, y) = \frac{1}{a}\sin(ax)$$

Exemplo 2.26 (Problema de valor de contorno)**.** *Agora seja a mesma EDP do exemplo* (2.25)*, mas com as condições* $f(0,0) = 0$ *e* $f(\frac{\pi}{2a}, 0) = 3$*. Da primeira condição, temos*

$$f(x, y) = C \cdot \sin(ax)$$

e da segunda:

$$f(\frac{\pi}{2a}, 0) = C \cdot \sin\left(\frac{\pi}{2}\right) = C = 3$$

e portanto a solução desse problema de valor de contorno é

$$f(x, y) = 3 \cdot \sin(ax)$$

2.4.4 Separação de variáveis

Algumas EDPs possuem soluções que podem ser reescritas como produto de funções de uma única variável. Podemos supor a existência dessas soluções para transformar uma EDP em várias EDOs mais simples de resolver. Se tratando de EDPs homogêneas, a função nula é sempre uma solução desse tipo, e por isso sempre tentamos evitá-la quando possível.

Note que se

$$f(x, y) = X(x)Y(y)$$

é solução de um problema de valor inicial $f(x, a) = F(x)$ e $\dfrac{\partial f}{\partial y}(x, a) = G(x)$, então

$$X(x)Y(a) = F(x) \quad \text{e} \quad X(x)Y'(a) = G(x)$$

o que significa que $F(x)$ só pode diferir de $G(x)$ por um múltiplo e então não temos a liberdade de escolher simultaneamente F e G. Em problemas cuja solução possuem variáveis separadas, é comum, portanto, usar condições de contorno ao invés de valor inicial.

Exemplo 2.27 (Equação do Calor)**.** *O problema de um sólido unidimensional com distribuição heterogênea de temperatura e sem nenhuma fonte de calor pode ser modelado pela EDP*

$$\frac{\partial u}{\partial t} = \eta \frac{\partial^2 u}{\partial x^2} \tag{144}$$

onde η *é o coeficiente de difusão (que depende do sólido) e* $u(x,t)$ *é a temperatura no ponto* x *e instante* t*. Para este problema, estamos interessados no intervalo* $[0, L]$ *e vamos considerar as condições de contorno de Dirichlet,* $u(0,t) = 0$ *e* $u(L,t) = 0$*.*

Vamos supor que existe alguma solução não-trivial dessa EDP tal que

$$u(x,t) = X(x)T(t) \tag{145}$$

Substituindo (145) *em* (144)*, obtemos*

$$X(x)T'(t) = \eta X''(x)T(t)$$

Sabendo que $u(x,t)$ *não é a função nula, podemos dividir essa expressão por* $X(x)T(t)$*:*

$$\frac{T'(t)}{T(t)} = \eta \frac{X''(x)}{X(x)} \tag{146}$$

Note que a igualdade se mantém independentemente das escolhas de x *e* t*, o que significa que ambos os membros são funções constantes de mesmo valor. Vamos denotar essa constante de* σ *e portanto podemos extrair duas EDOs da equação* (146)*:*

$$\frac{T'(t)}{T(t)} = \sigma \quad e \quad \eta \frac{X''(x)}{X(x)} = \sigma$$

A primeira equação se torna a EDO

$$T'(t) = \sigma T(t)$$

cuja solução é

$$T(t) = Ce^{\sigma t}$$

E a segunda equação se torna

$$X''(x) = \frac{\sigma}{\eta} X(x) \tag{147}$$

As soluções dessa EDO dependem do valor de $\frac{\sigma}{\eta}$. Uma vez que a constante η é sempre positiva, então a arbitrariedade é da constante σ. É preciso verificar se existem soluções não-nulas para $\sigma > 0$, $\sigma = 0$ e $\sigma < 0$.

A EDO no primeiro caso pode ser reescrita como

$$X''(x) = k^2 X(x) \tag{148}$$

onde $k = \sqrt{\dfrac{\sigma}{\eta}}$.Supondo $X(x) = e^{\lambda x}$, então

$$\lambda^2 e^{kx} = k^2 e^{kx}$$

Pela independência linear de e^{kx}, podemos afirmar que

$$\lambda^2 = k^2$$

ou seja, $\lambda_0 = k$ e $\lambda_1 = -k$. A solução geral de (148) é

$$X(x) = Ae^{kx} + Be^{-kx}$$

onde A e B são constantes arbitrárias.

Devemos testar as condições de contorno $u(0,t) = 0$ e $u(L,t) = 0$. Uma vez que não queremos que $u(x,t) = X(x)T(t)$ seja a função nula, então $T(t)$ também não é a função nula e consequentemente as condições de contorno indicam que

$$X(0) = 0 \quad \text{e} \quad X(L) = 0$$

Portanto

$$0 = Ae^{k0} + Be^{-k0} \implies B = -A$$
$$X(x) = Ae^{kx} - Ae^{-kx} = 2A \cdot \left[\frac{e^{kx} - e^{-kx}}{2}\right] = 2A\sinh{(kx)}$$

Já a segunda condição implica

$$0 = 2A\sinh{(kL)}$$

Visto que não estamos procurando a solução nula, não podemos escolher $A = 0$. E por outro lado, a função seno hiperbólico só se anula na origem, o que implicaria que $L = 0$ ou $k = 0$. O primeiro caso é inviável, uma vez que estamos modelando um sólido de comprimento não nulo. O segundo caso implica $k = 0$ e então $\sigma = 0$, o que viola a hipótese. Isso significa que não existe solução não-trivial com variáveis separadas para $\sigma > 0$.

Agora suponha que $\sigma = 0$. Então a EDO (147) se torna

$$X''(x) = 0$$

Sua solução geral é dada pela expressão

$$X(x) = Ax + B$$

As condições de contorno $X(0) = 0$ e $X(L) = 0$ implicam

$$0 = A0 + B \implies B = 0$$

$$0 = AL \implies A = 0$$

ou seja, não existe solução não-trivial para $\sigma = 0$.

Finalmente, suponha que $\sigma < 0$. A EDO (147) pode ser reescrita como

$$X''(x) = k^2 X(x) \tag{149}$$

onde

$$k = \sqrt{\frac{\sigma}{\eta}} = \sqrt{\frac{-|\sigma|}{\eta}} = i\sqrt{\frac{|\sigma|}{\eta}}$$

Substituindo $X(x) = e^{\lambda x}$ em (149), obtemos

$$\lambda^2 e^{\lambda x} = k^2 e^{\lambda x}$$

Pela independência linear de $e^{\lambda x}$, podemos afirmar que

$$\lambda^2 = k^2$$

ou seja

$$\lambda_0 = k \quad \text{e} \quad \lambda_1 = -k$$

Então a solução geral de (149) é

$$X(x) = Ae^{kx} + Be^{-kx}$$

Agora denote $k_0 = \sqrt{\frac{|\sigma|}{\eta}}$, de maneira que $k = ik_0$. Então podemos reescrever a solução na forma

$$X(x) = Ae^{ik_0 x} + Be^{-ik_0 x}$$

A temperatura do sólido $u(x,t)$ em qualquer ponto e instante de tempo é um valor real. Portanto, para restringir o domínio da solução aos reais impomos

$$\overline{X(x)} = X(x) \quad \forall x$$

e então

$$\overline{A}e^{-ik_0 x} + \overline{B}e^{ik_0 x} = Ae^{ik_0 x} + Be^{-ik_0 x}$$

Pela independência linear de $e^{ik_0 x}$ e $e^{-ik_0 x}$, temos que

$$\overline{A} = B \quad \text{e} \quad \overline{B} = A$$

Suponha $A = a + bi$. Então

$$X(x) = (a + bi)e^{ik_0 x} + (a - bi)e^{-ik_0 x}$$

$$X(x) = a \cdot (e^{ik_0 x} + e^{-ik_0 x}) + bi \cdot (e^{ik_0 x} - e^{-ik_0 x})$$

$$X(x) = 2a \cdot \left[\frac{e^{ik_0 x} + e^{-ik_0 x}}{2}\right] + 2bi^2 \cdot \left[\frac{e^{ik_0 x} - e^{-ik_0 x}}{2i}\right]$$

$$X(x) = 2a \cdot \cos(k_0 x) - 2b \cdot \sin(k_0 x)$$

Aplicando as condições de contorno $X(0) = 0$ e $X(L) = 0$:

$$0 = 2a \cdot \cos(0) - 2b \cdot \sin(0) \implies a = 0$$

$$0 = -2b \cdot \sin(k_0 L)$$

Não queremos $b = 0$, pois implica na solução trivial. Então apenas nos resta

$$\sin(k_0 L) = 0$$

ou seja,

$$k_0 L = n\pi \quad \forall n \in \mathbb{Z} \quad n \neq 0$$

$$k_0 = \frac{n\pi}{L}$$

Substituindo $k_0 = \sqrt{\frac{|\sigma|}{\eta}}$, obtemos

$$|\sigma_n| = \eta \frac{n^2\pi^2}{L^2}$$

Lembando que $\sigma_n < 0$, os possíveis valores de σ em que podemos obter soluções de variáveis separadas são, portanto,

$$\sigma_n = -\eta \frac{n^2\pi^2}{L^2} \quad \forall n \in \mathbb{Z} \quad n \neq 0$$

Portanto as soluções de (149) são

$$X_n(x) = A \cdot \sin \frac{n\pi x}{L}$$

Voltando à EDO da variável tempo, a sua solução é

$$T_n(t) = C \cdot \exp\left(-\eta \frac{n^2\pi^2}{L^2} t\right)$$

Uma solução da EDP do calor com variáveis separadas é, portanto,

$$u_n(x,t) = A_n \cdot \exp\left(-\eta \frac{n^2\pi^2}{L^2} t\right) \cdot \sin \frac{n\pi x}{L}$$

onde n é um inteiro arbitrário.

Note que para cada n temos uma solução linearmente independente. Então a solução geral dessa EDP com variáveis separadas é

$$u(x,t) = \sum_{n=1}^{\infty} A_n \cdot \exp\left(-\eta \frac{n^2\pi^2}{L^2} t\right) \cdot \sin \frac{n\pi x}{L} \tag{150}$$

Quando $t = 0$, obtemos a expressão

$$u(x,0) = \sum_{n=1}^{\infty} A_n \cdot \sin \frac{n\pi x}{L} \tag{151}$$

Se considerarmos que o problema de calor no sólido inicializa em $t = 0$, então a função $u(x, 0)$ nos diz a distribuição inicial da temperatura no sólido. Repare que $u(x, 0)$ é arbitrário, uma vez que as constantes A_n também são arbitrárias. Por exemplo, se definirmos

$$u(x,0) = 5 \cdot \sin \frac{\pi x}{L}$$

então $A_1 = 5$ e $A_n = 0$ para todo $n \geq 2$.

Mas e se definirmos $u(x, 0) = f(x)$, onde $f(x)$ é uma função arbitrária que satisfaz $f(0) = f(L) = 0$, será que podemos encontrar uma sequência $\{A_n\}$ que satisfaça

$$f(x) = \sum_{n=1}^{\infty} A_n \cdot \sin \frac{n\pi x}{L} \tag{152}$$

para todo $x \in [0, L]$? A equação (152) é um caso particular da série de Fourier, que será visto na próxima seção.

2.5 Série de Fourier

Vimos no exemplo (1.3) que se $x \in [0, 2\pi]$ e

$$u_0(x) = 1 \qquad u_{2n-1}(x) = cos(nx) \quad u_{2n}(x) = sin(nx) \quad \text{, para } n = 1, 2, 3, ...$$

então $m \neq n$ implica

$$\int_0^{2\pi} u_n(x) u_m(x) dx = 0 \tag{153}$$

Agora façamos a substituição

$$w = \frac{x \cdot L}{\pi} - L \quad \text{ou} \quad x = \frac{w\pi}{L} + \pi$$

A equação (153) se torna

$$\frac{\pi}{L} \int_{-L}^{L} u_n \left(\frac{w\pi}{L} + \pi \right) u_m \left(\frac{w\pi}{L} + \pi \right) dw = 0 \tag{154}$$

Temos que $\forall w \in [-L, L]$:

$$u_0(x) = 1 \implies u_0 \left(\frac{w\pi}{L} + \pi \right) = 1$$

$$u_{2n-1}(x) = cos(nx) \implies u_{2n-1} \left(\frac{w\pi}{L} + \pi \right) = \cos\left(\frac{wn\pi}{L} + n\pi \right)$$

$$u_{2n}(x) = sin(nx) \implies u_{2n}\left(\frac{w\pi}{L} + \pi\right) = \sin\left(\frac{wn\pi}{L} + n\pi\right)$$

$$\cos\left(\frac{wn\pi}{L} + n\pi\right) = \cos\left(\frac{wn\pi}{L}\right)\cdot\cos(n\pi) - \sin\left(\frac{wn\pi}{L}\right)\cdot\sin(n\pi) = (-1)^n \cos\left(\frac{wn\pi}{L}\right)$$

$$\sin\left(\frac{wn\pi}{L} + n\pi\right) = \sin\left(\frac{wn\pi}{L}\right)\cdot\cos(n\pi) + \cos\left(\frac{wn\pi}{L}\right)\cdot\sin(n\pi) = (-1)^n \sin\left(\frac{wn\pi}{L}\right)$$

Agora vamos definir

$$v_0(w) = 1$$

$$v_{2n-1}(w) = (-1)^n \cos\left(\frac{wn\pi}{L}\right)$$

$$v_{2n}(w) = (-1)^n \sin\left(\frac{wn\pi}{L}\right)$$

para todo $w \in [-L, L]$ e $n \geq 1$. Portanto

$$u_n\left(\frac{w\pi}{L} + \pi\right) = v_n(w)$$

$$u_m\left(\frac{w\pi}{L} + \pi\right) = v_m(w)$$

E então a expressão (154) se torna

$$\frac{\pi}{L}\int_{-L}^{L} v_n(w)v_m(w)dw = 0 \tag{155}$$

se $n \neq m$.

Note que multiplicar qualquer um dos termos por -1 não altera a validade da equação (155). Isso significa que se o produto interno é definido pela equação (6) com $w(x) = 1$, $a = -L$ e $b = L$, então o conjunto de funções

$$G = \left\{1, \cos\left(\frac{w\pi}{L}\right), \sin\left(\frac{w\pi}{L}\right), \cos\left(\frac{2w\pi}{L}\right), \ldots\right\} \tag{156}$$

é ortogonal e, consequentemente, linearmente independente.

Por uma questão de conveniência, trocaremos a variável muda w por x. Então uma função real $f(x)$ que pertence ao espaço gerado por esse conjunto pode ser expressada na forma

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cdot \cos\left(\frac{n\pi x}{L}\right) + b_n \cdot \sin\left(\frac{n\pi x}{L}\right) \tag{157}$$

A expressão (157) é chamada de Série de Fourier, onde $x \in [-L, L]$.

Podemos definir as constantes arbitrárias a_0, a_n e b_n usando o teorema (1.12):

$$\frac{a_0}{2} = \frac{\langle 1, f(x)\rangle}{\langle 1, 1\rangle}$$

$$a_n = \frac{\left\langle \cos\left(\frac{n\pi x}{L}\right), f(x)\right\rangle}{\left\langle \cos\left(\frac{n\pi x}{L}\right), \cos\left(\frac{n\pi x}{L}\right)\right\rangle}$$

$$b_n = \frac{\left\langle \sin\left(\frac{n\pi x}{L}\right), f(x) \right\rangle}{\left\langle \sin\left(\frac{n\pi x}{L}\right), \sin\left(\frac{n\pi x}{L}\right) \right\rangle}$$

Sabendo que $\sin(2\pi n) = 0$ e que $\sin(n\pi) = 0$:

$$\langle 1, 1 \rangle = \int_{-L}^{L} \mathrm{d}x = 2L$$

$$\left\langle \cos\left(\frac{n\pi x}{L}\right), \cos\left(\frac{n\pi x}{L}\right) \right\rangle = \int_{-L}^{L} \cos\left(\frac{n\pi x}{L}\right)^2 \mathrm{d}x = L \cdot \frac{\sin(2\pi n)}{2\pi n} + L = L$$

$$\left\langle \sin\left(\frac{n\pi x}{L}\right), \sin\left(\frac{n\pi x}{L}\right) \right\rangle = \int_{-L}^{L} \sin\left(\frac{n\pi x}{L}\right)^2 \mathrm{d}x = L - L \cdot \frac{\sin(2\pi n)}{2\pi n} = L$$

Então podemos concluir que

$$a_0 = \frac{1}{L} \int_{-L}^{L} f(x)\, \mathrm{d}x \tag{158}$$

$$a_n = \frac{1}{L} \int_{-L}^{L} f(x) \cdot \cos\left(\frac{n\pi x}{L}\right) \mathrm{d}x \tag{159}$$

$$b_n = \frac{1}{L} \int_{-L}^{L} f(x) \cdot \sin\left(\frac{n\pi x}{L}\right) \mathrm{d}x \tag{160}$$

Sabemos que a função cosseno é par e que a função seno é ímpar. Se $f(x)$ for uma função par, então

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cdot \cos\left(\frac{n\pi x}{L}\right) + b_n \cdot \sin\left(\frac{n\pi x}{L}\right) =$$

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cdot \cos\left(\frac{n\pi x}{L}\right) - b_n \cdot \sin\left(\frac{n\pi x}{L}\right)$$

o que implica

$$b_n = 0 \quad \forall n \geq 1$$

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cdot \cos\left(\frac{n\pi x}{L}\right) \tag{161}$$

Analogamente, se $f(x)$ for ímpar, temos

$$f(x) = \sum_{n=1}^{\infty} b_n \cdot \sin\left(\frac{n\pi x}{L}\right) \tag{162}$$

Note que a equação (162) se assemelha à (152).

2.5.1 Teorema de Fourier

Vimos na seção (1.4.1) que o elemento de um subespaço H de dimensão k denotado pela equação

$$\text{proj}_H(f) = \sum_{n=1}^{k} \frac{\langle v_n, f\rangle}{\langle v_n, v_n\rangle} v_n$$

é o elemento $h \in H$ que minimiza a norma $\|f - h\|$, onde v_n são os elementos de uma base de H.

Se $f(x)$ não pertence ao espaço gerado por G em (156), podemos encontrar uma função de G que minimiza a norma $\|f(x) - g(x)\|$. Essa é a projeção da função f em G, ou seja, a série de Fourier gerada pela função f usando as equações (158), (159) e (160).

Uma condição suficiente para que as integrais dessas três equações existam é que $f(x)$ seja uma função seccionalmente contínua, isto é, $f(x)$ seja limitada no intervalo $[-L, L]$ e que não tenha nenhuma descontinuidade ou que tenham apenas em um número finito de pontos.

Exemplo 2.28. *Seja a função $f(x) = x^2$ definida no intervalo $[-\pi, \pi]$. Como f é par, $b_n = 0$ para todo $n \geq 1$. E também*

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} x^2 \, \mathrm{d}x = \frac{1}{3\pi} x^3 \Big|_{-\pi}^{\pi} = \frac{2\pi^2}{3}$$

$$\begin{aligned}
a_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} x^2 \cos(nx) \, \mathrm{d}x \\
&= \frac{1}{\pi} \cdot \left[\frac{x^2 \sin(nx)}{n} \Big|_{-\pi}^{\pi} - \int_{-\pi}^{\pi} \frac{2x \sin(nx)}{n} \, \mathrm{d}x \right] \\
&= \frac{1}{\pi} \cdot \left[\frac{2x \cos(nx)}{n^2} \Big|_{-\pi}^{\pi} - \int_{-\pi}^{\pi} \frac{2}{n^2} \cdot \cos(nx) \right] \\
&= \frac{1}{\pi} \cdot \left[\frac{2\pi}{n^2} \cos(\pi n) + \frac{2\pi}{n^2} \cos(\pi n) - \left(\frac{2}{n^3} \sin(nx) \right) \Big|_{-\pi}^{\pi} \right] \\
&= \frac{4}{n^2} \cdot (-1)^n
\end{aligned}$$

Portanto

$$\text{proj}_G(f(x)) = \frac{\pi^2}{3} + \sum_{n=1}^{\infty} (-1)^n \frac{4}{n^2} \cos(nx) \tag{163}$$

As somas parciais dessa série de Fourier podem ser vistas na figura (1).
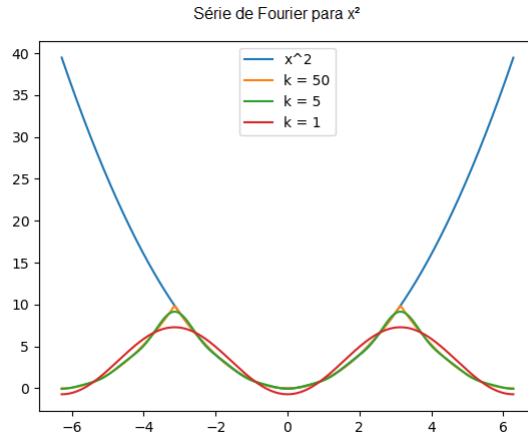
Figura 1: Somas parciais de $n = 1$ até k

Exemplo 2.29. *Seja f a função*

$$f(x) = \begin{cases} \frac{\pi}{2} & \text{se } x > 0 \\ 0 & \text{se } x = 0 \\ -\frac{\pi}{2} & \text{se } x < 0 \end{cases}$$

definida no intervalo $[-1, 1]$. Uma vez que $f(x)$ é ímpar, $a_0 = 0$ e $a_n = 0$ para todo $n \geq 1$. E também

$$\begin{aligned} b_n &= \int_{-1}^{1} f(x) \sin(n\pi x) \, \mathrm{d}x \\ &= \int_{-1}^{0} \frac{-\pi}{2} \sin(n\pi x) \, \mathrm{d}x + \int_{0}^{1} \frac{\pi}{2} \sin(n\pi x) \, \mathrm{d}x \end{aligned}$$

$$\int_{-1}^{0} -\frac{\pi}{2} \sin(n\pi x) \, \mathrm{d}x = \frac{\pi}{2} \cos(n\pi x) \cdot \frac{1}{n\pi} \bigg|_{-1}^{0} = \frac{1}{2n} \cdot (1 - \cos(n\pi))$$

$$\int_{0}^{1} \frac{\pi}{2} \sin(n\pi x) \, \mathrm{d}x = -\frac{\pi}{2} \cos(n\pi x) \cdot \frac{1}{n\pi} \bigg|_{0}^{1} = \frac{1}{2n} \cdot (1 - \cos(n\pi))$$

Sabendo que $\cos(n\pi) = (-1)^n$, então

$$b_n = \frac{1}{n} \cdot (1 - (-1)^n)$$

A constante b_n é zero quando n é par e é igual a $\frac{2}{n}$ quando é ímpar. Portanto

$$\mathrm{proj}_G(f(x)) = \sum_{n=1}^{\infty} \frac{2}{2n-1} \sin((2n-1)\pi x)$$

As somas parciais dessa série de Fourier podem ser vistas na figura (2).
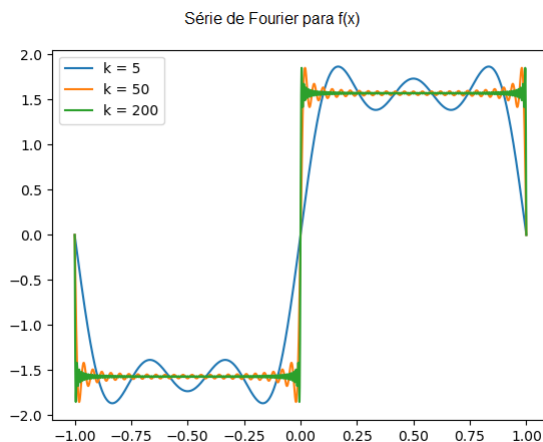
Figura 2: Somas parciais de $n = 1$ até k

Note que no exemplo (2.29), se alteramos o valor da função f na descontinuidade $x = 0$ para qualquer outro valor que não seja zero, f deixará de ser ímpar em qualquer subintervalo contendo a origem. Entretanto, as integrais das equações (158), (159) e (160) permanecem inalteradas quando f muda de valor em finitos pontos do intervalo $[a, b]$, o que significa que independentemente do valor que escolhermos para $f(0)$ a série de Fourier de f continuará convergindo para zero em $x = 0$, que é a média dos limites laterais de f em $x = 0$.

O próximo teorema nos diz as condições suficientes para que a igualdade

$$\text{proj}_G(f(x)) = f(x)$$

seja válida para todo $x \in \mathbb{R}$.

Teorema 2.33 (Teorema de Fourier)**.** *Seja f uma função periódica de período $2L$ seccionalmente diferenciável. Então a série de Fourier gerada por f converge pontualmente para a média dos seus limites laterais, ou seja*

$$\frac{f(x^+) + f(x^-)}{2} = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cdot \cos\Big(\frac{n\pi x}{L}\Big) + b_n \cdot \sin\Big(\frac{n\pi x}{L}\Big)$$

onde

$$f(x^+) = \lim_{h \to 0^+} f(x + h) \quad e \quad f(x^-) = \lim_{h \to 0^-} f(x + h)$$

Uma função f é seccionalmente diferenciável se sua derivada $f'(x)$ é seccionalmente contínua e se f também for seccionalmente contínua.

Demonstração. Vamos denotar o erro $E(x)$ entre a função f e sua série de Fourier como

$$E(x) = \text{proj}_G(f(x)) - \frac{1}{2}(f(x^+) + f(x^-))$$

A soma parcial $S_k(x)$ da série de Fourier é dada pela expressão

$$S_k(x) = \frac{a_0}{2} + \sum_{n=1}^{k} a_n \cdot \cos\left(\frac{n\pi x}{L}\right) + b_n \cdot \sin\left(\frac{n\pi x}{L}\right) \tag{164}$$

O erro parcial entre a função f e a série de Fourier gerada por ela é, portanto,

$$E_k(x) = S_k(x) - \frac{1}{2}(f(x^+) + f(x^-)) \tag{165}$$

Substituindo as equações (158), (159) e (160) em (164), obtemos

$$\begin{aligned} S_k(x) = \frac{1}{2L}\int_{-L}^{L} f(y)\,\mathrm{d}y + \\ \sum_{n=1}^{k}\left[\left(\frac{1}{L}\int_{-L}^{L} f(y)\cdot\cos\left(\frac{n\pi y}{L}\right)\mathrm{d}y\right)\cdot\cos\left(\frac{n\pi x}{L}\right) + \right. \\ \left.\left(\frac{1}{L}\int_{-L}^{L} f(y)\cdot\cos\left(\frac{n\pi y}{L}\right)\mathrm{d}y\right)\cdot\sin\left(\frac{n\pi x}{L}\right)\right] \end{aligned}$$

$$\begin{aligned} S_k(x) = \int_{-L}^{L}\left\{\frac{1}{2L} + \frac{1}{L}\sum_{n=1}^{k}\left[\cos\left(\frac{n\pi y}{L}\right)\cdot\cos\left(\frac{n\pi x}{L}\right) + \right.\right. \\ \left.\left.\sin\left(\frac{n\pi y}{L}\right)\cdot\sin\left(\frac{n\pi x}{L}\right)\right]\right\} f(y)\,\mathrm{d}y \end{aligned}$$

$$S_k(x) = \int_{-L}^{L}\left[\frac{1}{2L} + \frac{1}{L}\sum_{n=1}^{k}\cos\left(\frac{n\pi(x-y)}{L}\right)\right] f(y)\,\mathrm{d}y \tag{166}$$

Vamos denotar a expressão entre colchetes de núcleo de Dirichlet:

$$D_k(x) = \frac{1}{2L} + \frac{1}{L}\sum_{n=1}^{k}\cos\left(\frac{n\pi x}{L}\right) \tag{167}$$

e portanto podemos reescrever (166) como

$$S_k(x) = \int_{-L}^{L} D_k(x-y) f(y)\,\mathrm{d}y \tag{168}$$

O núcleo de Dirichlet possui as seguintes propriedades

1. $D_k(x)$ é uma função contínua, par e periódica de período $2L$.

2. $D_k(0) = \frac{1}{L}\left(k + \frac{1}{2}\right)$.

3. $\int_{-L}^{L} D_k(x)\,\mathrm{d}x = 1$

4. $D_k(x) = \frac{1}{2L} \cdot \frac{\sin\left[\left(k+\frac{1}{2}\right)\cdot\frac{\pi x}{L}\right]}{\sin\left(\frac{\pi x}{2L}\right)}$ se x não for um múltiplo inteiro de $2L$.

A demonstração dessas propriedades está no apêndice (A.1).

Fazendo a substituição $y = x - t$ em (168), obtemos

$$S_k(x) = \int_{x+L}^{x-L} D_k(t)\cdot f(x-t)(-\,\mathrm{d}t) = \int_{x-L}^{x+L} D_k(t)\cdot f(x-t)\,\mathrm{d}t$$

Uma vez que ambas as funções $D_k(t)$ e $f(x-t)$ são periódicas de período $2L$, o intervalo de integração pode ser transladado sem alterar seu valor. Então vamos fazer isso para o intervalo mais conveniente

$$S_k(x) = \int_{x-L}^{x+L} D_k(t)\cdot f(x-t)\,\mathrm{d}t = \int_{-L}^{L} D_k(t)\cdot f(x-t)\,\mathrm{d}t$$

Agora vamos dividir o intervalo de integração:

$$S_k(x) = \int_{-L}^{0} D_k(t)\cdot f(x-t)\,\mathrm{d}t + \int_{0}^{L} D_k(t)\cdot f(x-t)\,\mathrm{d}t$$

Na primeira integral, vamos fazer a substituição $t = -t'$:

$$S_k(x) = \int_{L}^{0} D_k(-t')\cdot f(x+t')(-\,\mathrm{d}t') + \int_{0}^{L} D_k(t)\cdot f(x-t)\,\mathrm{d}t$$

Sabemos que $D_k(t)$ é uma função par, e portanto $D_k(-t) = D_k(t)$. Por conveniência, voltamos a variável muda para t:

$$S_k(x) = \int_{0}^{L} D_k(t)\cdot f(x+t)\,\mathrm{d}t + \int_{0}^{L} D_k(t)\cdot f(x-t)\,\mathrm{d}t$$

$$S_k(x) = \int_{0}^{L} D_k(t)\cdot [f(x+t) + f(x-t)]\,\mathrm{d}t$$

A propriedade 3 do núcleo de Dirichlet nos diz que

$$\int_{-L}^{L} D_k(x)\,\mathrm{d}x = 1$$

Mas também sabemos que $D_k(x)$ é uma função par, e portanto

$$\int_{-L}^{L} D_k(x)\,\mathrm{d}x = 2\cdot\int_{0}^{L} D_k(x)\,\mathrm{d}x$$

e então

$$\int_0^L D_k(x)\,\mathrm{d}x = \frac{1}{2}$$

$$\frac{1}{2}(f(x^+) + f(x^-)) = \left(\int_0^L D_k(x)\,\mathrm{d}x\right)\cdot(f(x^+) + f(x^-))$$

E então podemos reescrever o erro parcial (165) como

$$E_k(x) = \int_0^L D_k(t)\cdot[f(x+t) + f(x-t)]\,\mathrm{d}t - \left(\int_0^L D_k(x)\,\mathrm{d}x\right)\cdot(f(x^+) + f(x^-))$$

$$E_k(x) = \int_0^L D_k(t)\cdot[f(x+t) + f(x-t)]\,\mathrm{d}t - \int_0^L D_k(t)\cdot(f(x^+) + f(x^-))\,\mathrm{d}t$$

$$E_k(x) = \int_0^L D_k(t)\cdot\left[(f(x+t) - f(x^+)) + (f(x-t) - f(x^-))\right]\mathrm{d}t$$

Vamos denotar a expressão entre colchetes de $g(x,t)$:

$$g(x,t) = (f(x+t) - f(x^+)) + (f(x-t) - f(x^-))$$

e então

$$E_k(x) = \int_0^L D_k(t)\cdot g(x,t)\,\mathrm{d}t$$

Vamos dividir o intervalo de integração novamente, onde $\delta > 0$:

$$E_k(x) = \int_0^\delta D_k(t)\cdot g(x,t)\,\mathrm{d}t + \int_\delta^L D_k(t)\cdot g(x,t)\,\mathrm{d}t$$

$$E_k(x) = \int_0^\delta t\cdot D_k(t)\cdot\frac{g(x,t)}{t}\,\mathrm{d}t + \int_\delta^L g(x,t)\cdot\frac{1}{2L}\cdot\frac{\sin\left[\left(k+\frac{1}{2}\right)\cdot\frac{\pi t}{L}\right]}{\sin\left(\frac{\pi t}{2L}\right)}\,\mathrm{d}t \quad (169)$$

Agora note que

$$\begin{aligned}
|t\cdot D_k(t)| &= \left|\sin\left((k+\frac{1}{2})\cdot\frac{t\pi}{L}\right)\cdot\frac{t}{2L\cdot\sin\left(\frac{\pi t}{2L}\right)}\right| \\
&\leq \frac{1}{2L}\cdot\left|\frac{t}{\sin\left(\frac{\pi t}{2L}\right)}\right| \quad \text{uma vez que} \quad 0 \leq \left|\sin\left((k+\frac{1}{2})\cdot\frac{t\pi}{L}\right)\right| \leq 1 \\
&\leq \frac{1}{2L}\cdot\left|\frac{L}{\sin\left(\frac{\pi L}{2L}\right)}\right| = \frac{1}{2L}\cdot|L| = \frac{1}{2} \quad \text{em} \quad t = L
\end{aligned}$$

$$|t \cdot D_k(t)| \leq \frac{1}{2}$$

E portanto

$$\left|\int_0^\delta t \cdot D_k(t) \cdot \frac{g(x,t)}{t}\,\mathrm{d}t\right| \leq \int_0^\delta |t \cdot D_k(t)| \cdot \left|\frac{g(x,t)}{t}\right| \mathrm{d}t \leq \frac{1}{2} \cdot \int_0^\delta \left|\frac{g(x,t)}{t}\right| \mathrm{d}t$$

Reescrevendo a função $\dfrac{g(x,t)}{t}$:

$$\frac{g(x,t)}{t} = \frac{f(x+t) - f(x^+)}{t} + \frac{f(x-t) - f(x^-)}{t}$$

Para que a integral de $\dfrac{g(x,t)}{t}$ seja finita em $(0,\delta]$, é necessário que o limite $\lim\limits_{t\to 0^+} \dfrac{g(x,t)}{t}$ exista e seja finito.

Por outro lado,

$$\lim_{t\to 0^+} \frac{g(x,t)}{t} = \lim_{t\to 0^+} \frac{f(x+t) - f(x^+)}{t} + \lim_{t\to 0^+} \frac{f(x-t) - f(x^-)}{t}$$

Mas por hipótese, $f(x)$ é seccionalmente diferenciável e portanto o limite existe e é igual a

$$\lim_{t\to 0^+} \frac{g(x,t)}{t} = f'(x^+) + f'(x^-)$$

Isso significa que $\dfrac{g(x,t)}{t}$ é limitada e então $\left|\dfrac{g(x,t)}{t}\right| \leq M$ para algum $M \in \mathbb{R}$:

$$\left|\int_0^\delta t \cdot D_k(t) \cdot \frac{g(x,t)}{t}\,\mathrm{d}t\right| \leq \left|\int_0^\delta M\,\mathrm{d}t\right| = \frac{M \cdot \delta}{2}$$

Escolhemos um δ tal que $\delta < \dfrac{\epsilon}{M}$ e consequentemente

$$\left|\int_0^\delta t \cdot D_k(t) \cdot \frac{g(x,t)}{t}\,\mathrm{d}t\right| \leq \frac{M \cdot \delta}{2} < \frac{M \cdot \epsilon}{2 \cdot M} = \frac{\epsilon}{2}$$

Isso significa que a integral tende a zero quando escolhemos um δ pequeno o suficiente.

Com o valor de $\delta > 0$ fixado, ainda sobra a segunda integral:

$$\int_\delta^L g(x,t) \cdot \frac{1}{2L} \cdot \frac{\sin\left[\left(k + \frac{1}{2}\right) \cdot \frac{\pi t}{L}\right]}{\sin\left(\frac{\pi t}{2L}\right)}\,\mathrm{d}t$$

O Lema de Riemann-Lebesgue (cuja demonstração está no apêndice (B)) nos diz que se f é uma função integrável e integrável em módulo em um intervalo $[a,b]$, então

$$\lim_{k\to\infty}\int_a^b f(x)\cdot\sin(kx)\,\mathrm{d}x = 0 \quad \text{e} \quad \lim_{k\to\infty}\int_a^b f(x)\cdot\cos(kx)\,\mathrm{d}x = 0$$

Como $\delta > 0$, o denominador da função

$$\frac{g(x,t)}{2L\cdot\sin\left(\dfrac{\pi t}{2L}\right)} \tag{170}$$

nunca se anula no intervalo $[\delta, L]$. Uma vez que $f(x)$ é seccionalmente contínua, então $g(x,t)$ é integrável em $[\delta, L]$ e portanto (170) também o é e satisfaz os requisitos do Lema de Riemann-Lebesgue. Isso significa que para um $\epsilon > 0$ qualquer, existe k_1 tal que

$$\left|\int_\delta^L D_k(t)\cdot g(x,t)\,\mathrm{d}t\right| < \frac{\epsilon}{2} \quad \text{sempre que} \quad k > k_1$$

ou seja

$$\begin{aligned}|E_k(x)| &= \left|\int_0^\delta D_k(t)\cdot g(x,t)\,\mathrm{d}t + \int_\delta^L D_k(t)\cdot g(x,t)\,\mathrm{d}t\right| \\ &\le \left|\int_0^\delta D_k(t)\cdot g(x,t)\,\mathrm{d}t\right| + \left|\int_\delta^L D_k(t)\cdot g(x,t)\,\mathrm{d}t\right| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon\end{aligned}$$

Em outras palavras,

$$\lim_{k\to\infty} E_k(x) = 0$$

ou então

$$\frac{f(x^+) + f(x^-)}{2} = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n\cdot\cos\left(\frac{n\pi x}{L}\right) + b_n\cdot\sin\left(\frac{n\pi x}{L}\right)$$

■

Exemplo 2.30. *Conforme calculamos no exemplo* (2.28)*, temos a expressão*

$$\mathrm{proj}_G(f(x)) = \frac{\pi^2}{3} + \sum_{n=1}^{\infty}(-1)^n\frac{4}{n^2}\cos(nx)$$

onde $f(x)$ é a parábola x^2 definida no intervalo $[-\pi,\pi]$. Agora suponha que $f(x)$ seja expandida na seguinte forma: $f(x)$ assume valores em todo o $\mathbb{R}$, $f(x) = x^2$

no intervalo $[-\pi, \pi]$ *e é periódica de período* 2π*. Como* $f(x)$ *é contínua, então no intervalo* $[-\pi, \pi]$

$$\frac{1}{2}(f(x^+) + f(x^-)) = f(x) = x^2$$

A função $f(x)$ *também é seccionalmente diferenciável, ou seja, podemos usar o teorema de Fourier para provar que no intervalo* $[-\pi, \pi]$ *a igualdade*

$$\text{proj}_G(f(x)) = x^2$$

é válida, e portanto

$$x^2 = \frac{\pi^2}{3} + \sum_{n=1}^{\infty}(-1)^n \frac{4}{n^2}\cos(nx) \quad \textit{se} \quad x \in [-\pi, \pi]$$

No caso particular em que $x = \pi$*, temos*

$$\pi^2 = \frac{\pi^2}{3} + \sum_{n=1}^{\infty}(-1)^n \frac{4}{n^2} \cdot (-1)^n$$

$$\sum_{n=1}^{\infty} \frac{4}{n^2} = \frac{2\pi^2}{3}$$

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$$

Note que a expressão do primeiro membro é a função Zeta de Riemann em $s = 2$*, ou seja*

$$\zeta(2) = \frac{\pi^2}{6}$$

2.5.2 Convergência uniforme da Série de Fourier

Dada a série de Fourier

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cdot \cos\left(\frac{n\pi x}{L}\right) + b_n \cdot \sin\left(\frac{n\pi x}{L}\right)$$

podemos afirmar que

$$\left|\sum_{n=1}^{\infty} a_n \cdot \cos\left(\frac{n\pi x}{L}\right) + b_n \cdot \sin\left(\frac{n\pi x}{L}\right)\right| \leq \sum_{n=1}^{\infty} \left|a_n \cdot \cos\left(\frac{n\pi x}{L}\right)\right| + \left|b_n \cdot \sin\left(\frac{n\pi x}{L}\right)\right|$$

Visto que o valor absoluto do seno e cosseno varia entre 0 e 1, então

$$\left|\sum_{n=1}^{\infty} a_n \cdot \cos\left(\frac{n\pi x}{L}\right) + b_n \cdot \sin\left(\frac{n\pi x}{L}\right)\right| \leq \sum_{n=1}^{\infty} |a_n| + |b_n|$$

Suponha que a série de Fourier convirja pontualmente para uma função f em um intervalo I. Se a_n e b_n forem coeficientes tais que a série

$$\sum_{n=1}^{\infty} |a_n| + |b_n| \tag{171}$$

converge, então pelo teorema de Weierstrass (2.25) a série de Fourier converge uniformemente para f no intervalo I.

Suponha que uma função f periódica de período $2L$ seja derivável, integrável e absolutamente integrável. Das equações (159) e (160), temos que

$$La_n = \int_{-L}^{L} f(x) \cdot \cos\Big(\frac{n\pi x}{L}\Big)\,\mathrm{d}x \tag{172}$$

$$Lb_n = \int_{-L}^{L} f(x) \cdot \sin\Big(\frac{n\pi x}{L}\Big)\,\mathrm{d}x \tag{173}$$

Agora vamos integrar por partes de modo que f seja derivada e o seno ou cosseno integrado. Começando pela equação (172):

$$La_n = \frac{L}{n\pi} f(x) \sin\Big(\frac{n\pi x}{L}\Big)\bigg|_{-L}^{L} - \frac{L}{n\pi}\int_{-L}^{L} f'(x) \sin\Big(\frac{n\pi x}{L}\Big)\,\mathrm{d}x$$

O primeiro termo do membro direito é zero e portanto

$$a_n = -\frac{1}{n\pi}\int_{-L}^{L} f'(x) \sin\Big(\frac{n\pi x}{L}\Big)\,\mathrm{d}x \tag{174}$$

Já para a equação (173):

$$Lb_n = -\frac{L}{n\pi} f(x) \cos\Big(\frac{n\pi x}{L}\Big)\bigg|_{-L}^{L} + \frac{L}{n\pi}\int_{-L}^{L} f'(x) \cos\Big(\frac{n\pi x}{L}\Big)\,\mathrm{d}x$$

O primeiro termo do segundo membro é zero devido a paridade do cosseno e então

$$b_n = \frac{1}{n\pi}\int_{-L}^{L} f'(x) \cos\Big(\frac{n\pi x}{L}\Big)\,\mathrm{d}x \tag{175}$$

Agora denote a'_n e b'_n os coeficientes da série de Fourier da função f'. Usando (174) e (175), obtemos

$$a_n = -\frac{L}{n\pi} b'_n \quad \text{e} \quad b_n = \frac{L}{n\pi} a'_n \tag{176}$$

Substituindo (176) na soma parcial de (171), temos que

$$\sum_{n=1}^{k} |a_n| + |b_n| = \frac{L}{\pi}\sum_{n=1}^{k} \frac{1}{n} \cdot (|a'_n| + |b'_n|)$$

E agora vamos usar o próximo teorema:

Teorema 2.34 (Desigualdade de Cauchy-Schwarz no $\mathbb{R}^n$)**.**

$$\left|\sum_{n=1}^{k} u_n v_n\right| \leq \left(\sum_{n=1}^{k} u_n^2\right)^{1/2} \left(\sum_{n=1}^{k} v_n^2\right)^{1/2} \tag{177}$$

Demonstração. Seja $\mathbb{R}^n$ o espaço euclidiano de dimensão n. A função produto escalar

$$\langle u, v\rangle = \sum_{k=1}^{n} u_n v_n \quad \forall u, v \in \mathbb{R}^n$$

obedece aos axiomas do produto interno e portanto podemos usar a desigualdade de Cauchy-Schwarz da álgebra linear:

$$|\langle u, v\rangle| \leq \|u\| \cdot \|v\|$$

que resulta em (177). ■

Fazendo $u_n = \dfrac{1}{n}$ e $v_n = (|a'_n| + |b'_n|)$ em (177), obtemos

$$\frac{L}{\pi}\sum_{n=1}^{k} \frac{1}{n} \cdot (|a'_n| + |b'_n|) \leq \frac{L}{\pi}\left(\sum_{n=1}^{k} \frac{1}{n^2}\right)^{1/2} \left(\sum_{n=1}^{k} (|a'_n| + |b'_n|)^2\right)^{1/2} \tag{178}$$

Agora note que para quaisquer a'_n e b'_n:

$$0 \leq (|a'_n| - |b'_n|)^2$$

$$0 \leq |a'_n|^2 - 2|a'_n b'_n| + |b'_n|^2$$

$$2|a'_n b'_n| \leq |a'_n|^2 + |b'_n|^2$$

Somando $|a'_n|^2 + |b'_n|^2$ em ambos os lados da equação, obtemos

$$|a'_n|^2 + 2|a'_n b'_n| + |b'_n|^2 \leq 2 \cdot (|a'_n|^2 + |b'_n|^2)$$

$$(|a'_n| + |b'_n|)^2 \leq 2 \cdot (|a'_n|^2 + |b'_n|^2)$$

Substituindo essa desigualdade em no segundo somatório de (178), obtemos a expressão

$$\sum_{n=1}^{k} |a_n| + |b_n| \leq \frac{\sqrt{2}L}{\pi}\left(\sum_{n=1}^{k} \frac{1}{n^2}\right)^{1/2} \left(\sum_{n=1}^{k} |a'_n|^2 + |b'_n|^2\right)^{1/2} \tag{179}$$

Quando $k \to \infty$, a série geométrica $\sum(\frac{1}{n})^2$ converge. Já para a segunda série de (179), uma condição suficiente para que ela convirja é dada pela desigualdade de Bessel

$$\frac{a_0^2}{2} + \sum_{n=1}^{\infty}\left(|a'_n|^2 + |b'_n|^2\right) \leq \frac{1}{L}\int_{-L}^{L} |f(x)|^2 \, \mathrm{d}x$$

de que f além de integrável seja quadrado integrável, ou seja, a integral de $|f|^2$ existe e é finita. A demonstração da desigualdade de Bessel será omitida, mas pode ser vista na seção 3.5 em [2].

Com isso, podemos enunciar o próximo teorema.

Teorema 2.35 (Primeiro teorema da convergência uniforme da série de Fourier)**.** *Seja f uma função periódica de período $2L$, contínua, quadrado integrável e com derivada primeira integrável. Então a série de Fourier de f converge uniformemente para f.*

Se f possuir descontinuidade em algum ponto x_0, então a série de Fourier não pode convergir uniformemente em qualquer intervalo que abranja x_0, uma vez que a série de Fourier é uma série de funções contínua em todo o $\mathbb{R}$. Entretanto, o próximo teorema nos diz que a série de Fourier de f pode convergir uniformemente para f em qualquer intervalo fechado em que f seja contínua.

Teorema 2.36 (Segundo teorema da convergência uniforme da série de Fourier)**.** *Seja f uma função periódica de período $2L$, seccionalmente contínua, quadrado integrável e com derivada primeira integrável. Então a série de Fourier de f converge uniformemente para f em qualquer intervalo fechado que não envolva pontos em que f seja descontínua.*

A demonstração desse teorema será omitida mas pode ser vista na seção 3.7 da referência [2].

Exemplo 2.31 (Continuação do exemplo (2.27))**.** *Vimos que a distribuição inicial de temperatura é dada pela série de Fourier*

$$u(x,0) = \sum_{n=1}^{\infty} A_n \cdot \sin \frac{n\pi x}{L}$$

Agora suponha que o valor de u em $t = 0$ é dado por $u(x,0) = f(x)$ no intervalo $[0, L]$:

$$f(x) = T_{max} \frac{4}{L^2} x(L-x) \quad \textit{se} \quad x \in [0, L]$$

A condição de Dirichlet $f(0) = f(L) = 0$ implica que a série de Fourier de f só deve ter senos, o que significa que f precisa ser uma função ímpar. Então vamos expandir f da seguinte forma:

$$f(x) = \begin{cases} T_{max} \frac{4}{L^2} x(L-x) & \textit{se} \quad x \in (0, L] \\ 0 & \textit{se } x = 0 \\ T_{max} \frac{4}{L^2} x(L+x) & \textit{se} \quad x \in [-L, 0) \\ \textit{Periódica de período } 2L \end{cases}$$

Dado que f é contínua, quadrado integrável e tem derivada primeira integrável, podemos usar o teorema (2.35) para afirmar que a série de Fourier de f vai

convergir uniformemente para f *em todo o* $\mathbb{R}$ *e*

$$A_n = \frac{1}{L}\int_{-L}^{L} f(x)\sin\Big(\frac{n\pi x}{L}\Big)\,\mathrm{d}x$$

Como $f(x)\sin\Big(\frac{n\pi x}{L}\Big)$ *é par, podemos calcular* A_n *na seguinte forma*

$$A_n = \frac{2}{L}\int_{0}^{L} f(x)\sin\Big(\frac{n\pi x}{L}\Big)\,\mathrm{d}x = T_{max}\frac{4}{L^2}\cdot\frac{2}{L}\int_{0}^{L} x(L-x)\sin\Big(\frac{n\pi x}{L}\Big)\,\mathrm{d}x$$

$$A_n = T_{max}\frac{4}{L^2}\frac{4L^2}{n^3\pi^3}(1-(-1)^n) = T_{max}\frac{16}{n^3\pi^3}(1-(-1)^n)$$

e portanto a solução geral do problema com variáveis separadas é

$$u(x,t) = T_{max}\frac{16}{\pi^3}\sum_{n=1}^{\infty}\frac{(1-(-1)^n)}{n^3}\cdot\exp\Big(-\eta\frac{n^2\pi^2}{L^2}t\Big)\cdot\sin\frac{n\pi x}{L}$$

As somas parciais de $u(x,0)$ *com* $T_{max} = 6$ *e* $L = 7$ *podem ser vistas na figura* (3) *enquanto que a função* $u(x,t)$ *com esses parâmetros no intervalo* $[0,L]$ *pode ser vista na figura* (4)
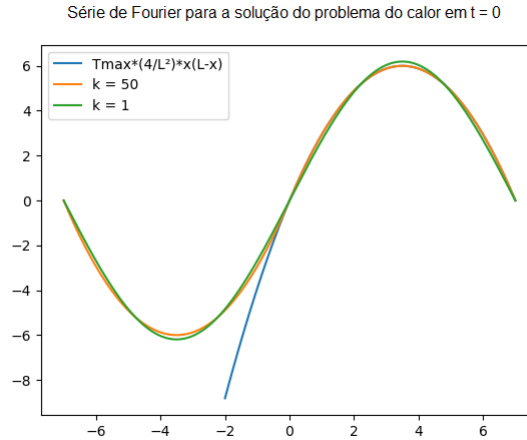
Figura 3: Somas parciais de $n = 1$ até k de $u(x,0)$

Note na figura (4) *que conforme o tempo passa, a temperatura na parte central do sólido decai lentamente até ficar igual a temperatura das pontas (que é fixa e pode ser considerado como uma temperatura ambiente).*
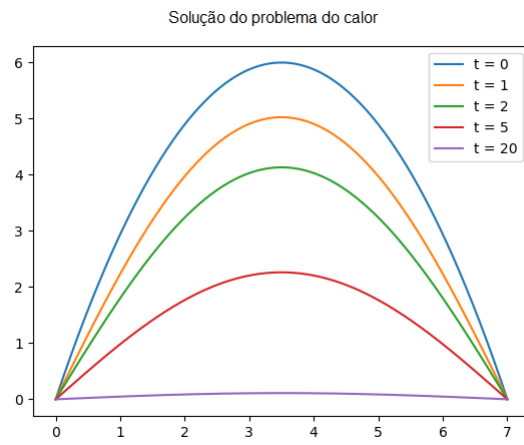
Figura 4: Somas parciais com $k = 50$ de $u(x, t)$ para diferentes valores de t

3 Introdução à Mecânica Quântica

3.1 A introdução da energia quantizada na física

No início do século 20, houve debates acerca do problema envolvendo o espectro de frequências emitidas por corpos negros (corpos que não refletem nenhuma luz incidida sobre eles) em que os resultados experimentais divergiam das previsões do modelo teórico da termodinâmica da época.

Um exemplo de corpo negro é um objeto oco com uma abertura pequena o suficiente para que qualquer radiação eletromagnética em seu interior reflita por todo o espaço interno até ser completamente absorvida pelo material, independentemente da sua composição. Dessa forma, o índice de absorção de luz é praticamente total e o índice de reflexão quase nulo, ou seja, não há retorno da luz que adentra o objeto. Apesar disso, o objeto ainda emite radiação por essa abertura, sendo essa concernente aos movimentos térmicos das moléculas que compõem o corpo. Podemos modelar esses movimentos por meio de um oscilador harmônico de frequência ν.

Dado um corpo negro desse tipo com uma temperatura T em Kelvin, a distribuição espectral da radiação térmica desse corpo é dada pela lei de Wien

$$\rho_\nu = \nu^3 \cdot f\left(\frac{\nu}{T}\right) \tag{180}$$

para alguma função f. No caso em que a energia é irradiada pela abertura na mesma proporção em que a energia é absorvida por ela, essa equação assume a forma

$$\rho_\nu = \frac{8\pi\nu^2}{c^3}\langle\epsilon\rangle \tag{181}$$

onde $\langle\epsilon\rangle$ é a energia média dos osciladores. Na mecânica clássica assume-se que em os valores de ν são contínuos, então nesse caso a energia média é dada por

$$\langle\epsilon\rangle = \int_0^\infty \epsilon p(\epsilon)\,\mathrm{d}\epsilon$$

A densidade $p(\epsilon)$ é dada pela distribuição de Boltzmann:

$$p(\epsilon) = \frac{1}{Z}\exp\left(-\frac{\epsilon}{K_BT}\right) \tag{182}$$

onde K_B é a constante de Boltzmann (aproximadamente $1.38 \cdot 10^{-23}J/K$) e Z é a constante de normalização tal que

$$\int_0^\infty p(\epsilon)\,\mathrm{d}\epsilon = 1$$

no que resulta

$$\frac{1}{Z}\int_0^\infty \exp\left(-\frac{\epsilon}{K_BT}\right)\mathrm{d}\epsilon = \frac{1}{Z}\cdot(-K_BT)\cdot\exp\left(-\frac{\epsilon}{K_BT}\right)\Bigg|_0^\infty = \frac{K_BT}{Z} = 1$$

$$Z = K_B T$$

E então

$$\begin{aligned}
\langle \epsilon \rangle &= \int_0^\infty \epsilon p(\epsilon)\,\mathrm{d}\epsilon \\
&= \frac{1}{K_B T} \int_0^\infty \epsilon \exp\left(-\frac{\epsilon}{K_B T}\right) \mathrm{d}\epsilon \\
&= \frac{1}{K_B T} \cdot \left[\underbrace{K_B T \cdot \epsilon \cdot \exp\left(-\frac{\epsilon}{K_B T}\right)\Bigg|_\infty^0}_{=0} + K_B T \int_0^\infty \exp\left(-\frac{\epsilon}{K_B T}\right) \mathrm{d}\epsilon \right] \\
&= \frac{1}{K_B T} \cdot (K_B T)^2 \\
&= K_B T
\end{aligned}$$

A equação (181) se torna, portanto:

$$\rho_\nu = \frac{8\pi\nu^2}{c^3} K_B T \tag{183}$$

Essa expressão (183) é a lei de Rayleigh-Jeans, uma previsão teórica da distribuição espectral de um corpo negro a uma temperatura T cujo comportamento é parabólico. Entretanto, os experimentos mostram que para uma certa frequência ν, a energia irradiada nesse espectro começa a diminuir drasticamente conforme ν aumenta, discrepando da parábola da equação (183), conforme pode ser visto na figura (5).

Para corrigir esse problema, Max Planck propôs restringir a liberdade dos valores de ν de forma que ϵ só assuma valores múltiplos de uma constante ϵ_ν, o quantum. Dessa forma, a densidade de energia é calculada usando a mesma expressão (182) com $\epsilon = n\epsilon_\nu$, mas o método de normalização da constante Z agora é discreto:

$$\sum_{n=0}^{\infty} p(\epsilon_\nu) = 1$$

ou seja

$$Z = \sum_{n=0}^{\infty} \exp\left(-\frac{n\epsilon_\nu}{K_B T}\right)$$

Note que a série infinita é uma série geométrica de razão $r = \exp\left(-\dfrac{\epsilon_\nu}{K_B T}\right)$. Como o expoente é negativo, então $r < 1$ e podemos usar a equação (89):

$$Z = \frac{1}{1 - \exp\left(-\dfrac{\epsilon_\nu}{K_B T}\right)}$$
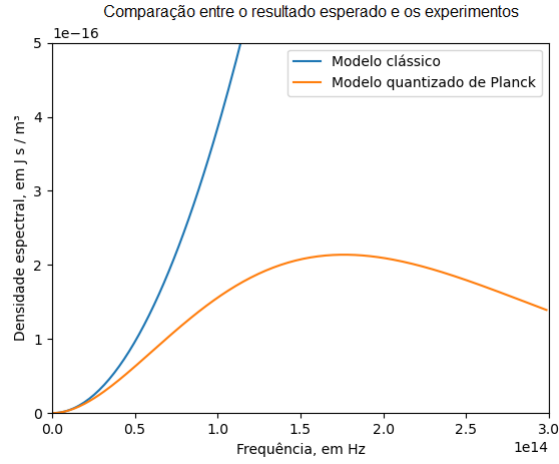
Figura 5: Comparação entre a distribuição espectral teórica da mecânica clássica com a obtida experimentalmente (posteriormente calculada por Planck) com $T = 3000$ Kelvin.

O cálculo da energia média também é discreto:

$$\begin{aligned}\langle \epsilon \rangle &= \sum_{n=0}^{\infty} n\epsilon_\nu p(\epsilon_\nu) \\ &= \sum_{n=0}^{\infty} \frac{1}{Z} \cdot n\epsilon_\nu \exp\left(-\frac{n\epsilon_\nu}{K_B T}\right) \\ &= \frac{1}{Z} \cdot \epsilon_\nu \sum_{n=0}^{\infty} n \cdot \exp\left(-\frac{n\epsilon_\nu}{K_B T}\right)\end{aligned}$$

Para calcularmos o somatório, fazemos $r = \exp\left(-\frac{\epsilon_\nu}{K_B T}\right)$ e notar que

$$\begin{aligned}\sum_{n=0}^{\infty} n r^n &= r \cdot \frac{d}{dr} \sum_{n=0}^{\infty} r^n \\ &= r \cdot \frac{d}{dr} \frac{1}{1-r} \\ &= r \cdot \frac{1}{(1-r)^2}\end{aligned}$$

e então

$$\langle\epsilon\rangle = \frac{1}{Z}\cdot\epsilon_\nu\cdot\frac{\exp\left(-\frac{\epsilon_\nu}{K_BT}\right)}{\left(1-\exp\left(-\frac{\epsilon_\nu}{K_BT}\right)\right)^2}$$

$$= \frac{\epsilon_\nu\cdot\exp\left(-\frac{\epsilon_\nu}{K_BT}\right)}{1-\exp\left(-\frac{\epsilon_\nu}{K_BT}\right)}$$

Multiplicando o numerador e o denominador por $\exp\left(\frac{\epsilon_\nu}{K_BT}\right)$, obtemos

$$\langle\epsilon\rangle = \frac{\epsilon_\nu}{\exp\left(\frac{\epsilon_\nu}{K_BT}\right)-1}$$

Portanto a equação (181) se torna

$$\rho_\nu = \frac{8\pi\nu^2}{c^3}\cdot\frac{\epsilon_\nu}{\exp\left(\frac{\epsilon_\nu}{K_BT}\right)-1} \tag{184}$$

Para que essa equação tome a forma da lei de Wien (180), é necessário que ϵ_ν seja algum múltiplo de ν, ou seja, $\epsilon_\nu = h\cdot\nu$, o que resulta

$$\rho_\nu = \frac{8\pi h\nu^3}{c^3}\cdot\frac{1}{\exp\left(\frac{h\cdot\nu}{K_BT}\right)-1} \tag{185}$$

cujo gráfico está na figura (5). Fazendo a aproximação de Taylor de primeira ordem na exponencial de (185) com $h\nu$ muito menor do que K_BT e desprezando os termos de segunda ordem ou mais, obtemos

$$\rho_\nu \approx \frac{8\pi h\nu^3}{c^3}\cdot\frac{1}{\frac{h\cdot\nu}{K_BT}} = \frac{8\pi\nu^2}{c^3}K_BT$$

e então recuperamos o resultado clássico (183) para frequências baixas.

3.1.1 O efeito fotoelétrico

Os experimentos da radiação do corpo negro apontam para uma quantização da radiação emitida pelos osciladores térmicos. Por outro lado, um outro experimento semelhantemente apontava para uma recepção discreta da energia da luz pelos elétrons. Essa ideia foi proposta por Einstein para explicar o porquê dos elétrons ejetados de uma placa metálica não ganharem energia cinética ao ampliar a intensidade da luz que a incide. Podemos descrever o fenômeno como uma transferência de energia do fóton $h\nu$ para um elétron, cuja energia necessária para se libertar do átomo é dada por Φ e a energia restante é convertida em cinética K. Na forma de equação, temos

$$h\nu = \Phi + K \tag{186}$$

A energia cinética do elétron K também pode ser reescrita como $K = qV$, onde q é a carga do elétron e V é a diferença de potencial acarretada pelos elétrons:

$$h\nu = \Phi + qV$$

Obtendo os valores de ν e V experimentalmente, foi possível calcular a constante de Planck:

$$h \approx 6.626 \cdot 10^{-34} J \cdot s$$

3.2 A Equação de Schrödinger

Demonstrado que uma radiação eletromagnética pode ser tratada como uma partícula em certos casos, De Broglie fez uma hipótese de que partículas (como elétrons) podem também se comportar como ondas, o que foi provado anos depois. Nesse contexto, Schrödinger formulou uma equação para determinar a função que representa essa onda de matéria, equação que foi baseada na formulação da mecânica clássica de Hamilton-Jacobi, que melhor se adéqua aos problemas ondulatórios.

A equação diferencial parcial de Schrödinger (188) foi formulada de uma forma que se os níveis de energia trabalhados forem muito maiores do que o quantum (ou similarmente, $h \to 0$), a EDP retorna à equação da mecânica clássica de Hamilton-Jacobi

$$\frac{1}{2m}\left\|\vec{\nabla} S\right\|^2 + V + \frac{\partial S}{\partial t} = 0 \tag{187}$$

cujos detalhes podem ser vistos em [3].

Na mecânica quântica, a função de onda $\psi(x,t)$ é suficiente para descrever qualquer partícula, e se conhecermos todas as funções de onda de todas as partículas, conhecemos todo o sistema físico. A evolução temporal da função de onda é regida pela EDP de Schrödinger (188)

$$i\hbar\frac{\partial \psi}{\partial t}(x,t) = -\frac{\hbar^2}{2m}\frac{\partial^2 \psi}{\partial x^2}(x,t) + U(x,t)\psi(x,t) \tag{188}$$

$$\text{onde} \quad \hbar = \frac{h}{2\pi}$$

e portanto a mecânica quântica é determinística quanto ao estado $\psi(x,t)$ de cada partícula.

Suponha um sistema físico envolvendo uma partícula, como o elétron, em um eixo unidimensional x. A função de onda, embora em geral seja complexa, fornece todas as informações possíveis que podem ser extraídas dessa partícula. Na interpretação de Copenhague, a função

$$\rho(x,t) = |\psi(x,t)|^2 = \psi(x,t)^*\psi(x,t)$$

nos diz a densidade de probabilidade de encontrarmos tal partícula no ponto x e no instante t caso haja alguma medição para tentar localizar a sua posição.

Aqui há outra diferença da mecânica clássica: as partículas não possuem posições bem definidas. O princípio da incerteza de Heisenberg nos diz que se σ_x é o desvio padrão dos resultados experimentais para tentar encontrar a posição da partícula e σ_p é o desvio padrão dos resultados para tentar encontrar o momento da partícula, então

$$\sigma_x \cdot \sigma_p \geq \frac{\hbar}{2} \tag{189}$$

Isso significa que nenhum dos desvios padrões pode ser zero, ou em outras palavras nem a posição nem o momento da partícula pode ser determinado com total precisão. E se aumentarmos a precisão da posição, aumentamos a incerteza do momento e vice-versa.

3.2.1 Normalização da função de onda

Uma vez que a equação de Schrödinger é uma EDP linear homogênea, se $\psi(x,t)$ é uma solução então $B \cdot \psi(x,t)$ também é. Mas $\rho(x,t) = |\psi(x,t)|^2$ também é uma função densidade de probabilidade. Logo $|\psi(x,t)|^2$ precisa satisfazer

$$\int_{-\infty}^{\infty} |\psi(x,t)|^2 \, \mathrm{d}x = 1 \tag{190}$$

para todo t. Entretanto, há soluções da equação de Schrödinger cuja integral (190) diverge, e então é impossível escolher a constante B para que satisfaça (190). Porém, se $\psi(x,t)$ for quadrado integrável, então a integral impropria da densidade será finita e poderemos escolher B que normaliza a solução.

Todavia ainda há a hipótese de $\psi(x,t)$ ser normalizada para um tempo $t = t_0$ mas não ser normalizada para algum outro t. Entretanto o próximo teorema rejeita essa possibilidade se $\psi(x,t)$ for quadrado integrável.

Definição 3.1. *Uma solução da equação de Schrödinger é dita normalizável se* (190) *for válida para todo t.*

Teorema 3.1. *Se $\psi(x,t)$ é uma solução da equação de Schrödinger e é quadrado integrável, então $\psi(x,t)$ é normalizável.*

Demonstração. Nessas condições, precisamos provar que (190) é invariante em relação ao tempo, ou seja:

$$\frac{d}{dt}\left[\int_{-\infty}^{\infty} |\psi(x,t)|^2 \, \mathrm{d}x\right] = 0$$

Usando a fórmula de Leibniz, obtemos uma expressão equivalente:

$$\int_{-\infty}^{\infty} \frac{\partial}{\partial t} |\psi(x,t)|^2 \, \mathrm{d}x = 0$$

Usando a definição de módulo e a regra do produto de derivadas:

$$\frac{\partial}{\partial t}|\psi(x,t)|^2 = \frac{\partial}{\partial t}\psi(x,t)^*\psi(x,t) \tag{191}$$

$$= \frac{\partial}{\partial t}\left(\psi(x,t)^*\right)\psi(x,t) + \frac{\partial}{\partial t}\left(\psi(x,t)\right)\psi(x,t)^* \tag{192}$$

Agora isolando o termo $\dfrac{\partial \psi}{\partial t}$ na equação (188), obtemos

$$\frac{\partial \psi}{\partial t} = i\frac{\hbar}{2m}\frac{\partial^2 \psi}{\partial x^2}(x,t) - \frac{i}{\hbar}U(x,t)\psi(x,t) \tag{193}$$

E para o conjugado dessa equação também temos

$$\left(\frac{\partial \psi}{\partial t}\right)^* = -i\frac{\hbar}{2m}\frac{\partial^2 \psi^*}{\partial x^2}(x,t) + \frac{i}{\hbar}U(x,t)\psi(x,t)^* \tag{194}$$

Substituindo (193) e (194) em (192), obtemos

$$\frac{\partial}{\partial t}|\psi(x,t)|^2 = -i\frac{\hbar}{2m}\frac{\partial^2 \psi^*}{\partial x^2}(x,t)\psi(x,t) + \frac{i}{\hbar}U(x,t)\psi(x,t)^*\psi(x,t) + i\frac{\hbar}{2m}\frac{\partial^2 \psi}{\partial x^2}(x,t)\psi(x,t)^* - \frac{i}{\hbar}U(x,t)\psi(x,t)\psi(x,t)^*$$

$$\frac{\partial}{\partial t}|\psi(x,t)|^2 = \frac{i\hbar}{2m}\cdot\left[\frac{\partial^2 \psi}{\partial x^2}(x,t)\psi(x,t)^* - \frac{\partial^2 \psi^*}{\partial x^2}(x,t)\psi(x,t)\right]$$

Agora vamos aplicar simplificar o termo em parênteses da seguinte forma: se f e g são funções duas vezes diferenciáveis, então

$$\frac{d}{dx}(f'g - fg') = f''g - fg'' \tag{195}$$

Demonstração. Aplicando a regra do produto, obtemos

$$\frac{d}{dx}(f'g - fg') = f''g + f'g' - f'g' - fg'' = f''g - fg''$$

■

Usando $f = \psi(x,t)^*$ e $g = \psi(x,t)$ em (195), chegamos à expressão

$$\frac{\partial}{\partial t}|\psi(x,t)|^2 = \frac{i\hbar}{2m}\cdot\frac{\partial}{\partial x}\left[\frac{\partial \psi}{\partial x}(x,t)\psi(x,t)^* - \frac{\partial \psi^*}{\partial x}(x,t)\psi(x,t)\right] \tag{196}$$

e então

$$\int_{-\infty}^{\infty}\frac{\partial}{\partial t}|\psi(x,t)|^2\,\mathrm{d}x = \int_{-\infty}^{\infty}\frac{i\hbar}{2m}\cdot\frac{\partial}{\partial x}\left[\frac{\partial \psi}{\partial x}(x,t)\psi(x,t)^* - \frac{\partial \psi^*}{\partial x}(x,t)\psi(x,t)\right]\mathrm{d}x$$

Como $\psi(x,t)$ é quadrado integrável, suas derivadas em x também o são, e podemos usar o teorema fundamental do cálculo para concluir que

$$\int_{-\infty}^{\infty} \frac{\partial}{\partial t}|\psi(x,t)|^2 \, \mathrm{d}x = \frac{i\hbar}{2m} \cdot \left[\frac{\partial \psi}{\partial x}(x,t)\psi(x,t)^* - \frac{\partial \psi^*}{\partial x}(x,t)\psi(x,t) \right] \Bigg|_{-\infty}^{\infty} \tag{197}$$

Uma vez que a função densidade $\psi(x,t)^*\psi(x,t)$ obedece (190), que é uma integral imprópria convergente de $-\infty$ a ∞, então obrigatoriamente

$$\lim_{x\to\infty} |\psi(x,t)|^2 = 0 \quad \text{e} \quad \lim_{x\to-\infty} |\psi(x,t)|^2 = 0 \tag{198}$$

Isso também é válido para $\psi(x,t)$ e $\dfrac{\partial \psi}{\partial x}(x,t)$, o que implica

$$\int_{-\infty}^{\infty} \frac{\partial}{\partial t}|\psi(x,t)|^2 \, \mathrm{d}x = 0$$

Ou seja, se normalizarmos a função de onda $\psi(x,t)$ para algum $t = t_0$, então ψ continuará normalizada para qualquer outro t. ■

Note que embora a densidade de probabilidade possa mudar com o tempo, a integral dessa densidade será sempre 1.

3.2.2 Momento linear

Dada a densidade de probabilidade $|\psi(x,t)|^2$ dos resultados de medida da posição de um certo estado $\psi(x,t)$, o valor esperado da posição é dado pela expressão

$$\langle x \rangle = \int_{-\infty}^{\infty} x \cdot |\psi(x,t)|^2 \, \mathrm{d}x$$

Geralmente o valor de $\langle x \rangle$ pode mudar com o tempo (o que não é o caso quando o potencial não depende do tempo, conforme veremos adiante) e assim podemos calcular a velocidade com a qual o valor esperado $\langle x \rangle$ se move no eixo x:

$$\frac{d\langle x \rangle}{dt} = \frac{d}{dt}\int_{-\infty}^{\infty} x \cdot |\psi(x,t)|^2 \, \mathrm{d}x = \int_{-\infty}^{\infty} x \cdot \frac{\partial}{\partial t}|\psi(x,t)|^2 \, \mathrm{d}x \tag{199}$$

Podemos usar a equação (196) e substituir em (199):

$$\frac{d\langle x \rangle}{dt} = \frac{i\hbar}{2m}\int_{-\infty}^{\infty} x \cdot \frac{\partial}{\partial x}\left[\frac{\partial \psi}{\partial x}\psi^* - \frac{\partial \psi^*}{\partial x}\psi\right] \mathrm{d}x$$

Agora vamos integrar por partes, derivando x e integrando o termo com a derivada parcial:

$$\frac{d\langle x \rangle}{dt} = \frac{i\hbar}{2m} \cdot x \cdot \left[\frac{\partial \psi}{\partial x}\psi^* - \frac{\partial \psi^*}{\partial x}\psi\right] \Bigg|_{-\infty}^{\infty} - \frac{i\hbar}{2m}\int_{-\infty}^{\infty} \left[\frac{\partial \psi}{\partial x}\psi^* - \frac{\partial \psi^*}{\partial x}\psi\right] \mathrm{d}x$$

Considerando os limites (198), o primeiro termo tende a zero (estamos supondo que $\psi(x,t)$ tende a zero mais rapidamente do que x tende a infinito).

$$\frac{d\langle x\rangle}{dt} = -\frac{i\hbar}{2m}\int_{-\infty}^{\infty}\left[\frac{\partial\psi}{\partial x}\psi^* - \frac{\partial\psi^*}{\partial x}\psi\right]\mathrm{d}x \tag{200}$$

Vamos usar integração por partes novamente em ambos os termos, integrando o termo que tem a derivada em x e derivando o restante:

$$\int_{-\infty}^{\infty}\frac{\partial\psi}{\partial x}\psi^*\,\mathrm{d}x = \psi\psi^*\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty}\frac{\partial\psi^*}{\partial x}\psi\,\mathrm{d}x = -\int_{-\infty}^{\infty}\frac{\partial\psi^*}{\partial x}\psi\,\mathrm{d}x$$

$$\int_{-\infty}^{\infty}\frac{\partial\psi^*}{\partial x}\psi\,\mathrm{d}x = \psi^*\psi\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty}\frac{\partial\psi}{\partial x}\psi^*\,\mathrm{d}x = -\int_{-\infty}^{\infty}\frac{\partial\psi}{\partial x}\psi^*\,\mathrm{d}x$$

Note que permutar o conjugado entre $\frac{\partial\psi}{\partial x}$ e ψ^* apenas troca o sinal da integral. Então

$$\frac{\partial\psi}{\partial x}\psi^* - \frac{\partial\psi^*}{\partial x}\psi = 2\cdot\left[\frac{\partial\psi}{\partial x}\psi^*\right]$$

Substituindo em (200), obtemos

$$\frac{d\langle x\rangle}{dt} = -\frac{i\hbar}{m}\int_{-\infty}^{\infty}\frac{\partial\psi}{\partial x}\psi^*\,\mathrm{d}x \tag{201}$$

A velocidade do valor esperado da posição da partícula é também o valor esperado das medidas de velocidade da partícula. A demonstração está no capítulo 3 de [4] usando a velocidade como o observável Q.

Desde o início estamos supondo a massa constante, então o análogo quântico do momento linear $p = mv$ é dado por

$$\langle p\rangle = m\langle v\rangle = -i\hbar\int_{-\infty}^{\infty}\frac{\partial\psi}{\partial x}\psi^*\,\mathrm{d}x = \frac{\hbar}{i}\int_{-\infty}^{\infty}\frac{\partial\psi}{\partial x}\psi^*\,\mathrm{d}x$$

Reescrevendo $\langle x\rangle$ e $\langle p\rangle$ de maneira mais conveniente:

$$\langle x\rangle = \int_{-\infty}^{\infty}\psi^*\cdot(x)\cdot\psi\,\mathrm{d}x$$

$$\langle p\rangle = \int_{-\infty}^{\infty}\psi^*\left(\frac{\hbar}{i}\frac{\partial}{\partial x}\right)\psi\,\mathrm{d}x$$

Isso nos dá pistas de que um observável pode ser calculado escolhendo um operador apropriado entre ψ^* e ψ. No caso da posição esperada é o operador multiplicador x, e o do momento esperado é o operador $\frac{\hbar}{i}\frac{\partial}{\partial x}$.

3.2.3 Função de onda estacionária

Na mecânica clássica, o potencial $U(x,t)$ fornece o tipo de experimento que está sendo realizado. Por exemplo se usarmos o potencial $\frac{kx^2}{2}$ na lei de movimento unidimensional de Newton

$$F = -\frac{dU}{dx} = \frac{dp}{dt}$$

no caso em que a massa é constante, ela irá retornar a equação diferencial do oscilador harmônico simples

$$\frac{d^2}{dt^2}x(t) = -\frac{k}{m}x(t)$$

cuja solução determina a posição da partícula. Derivando a solução obtemos a velocidade dela e consequentemente também determinamos o estado desse sistema em qualquer instante t.

Isso também é válido na mecânica quântica, pois dado um potencial $U(x,t)$ podemos obter a função de onda que descreve o sistema físico desse potencial resolvendo a equação diferencial de Schrödinger (188).

O caso mais simples é o de que o potencial só depende da posição, o que implica que podemos escrever $U(x,t) = U(x)$. Nessa condição, podemos usar a separação de variáveis para resolver a equação de Schrödinger:

$$\psi(x,t) = \Psi(x) \cdot \Phi(t)$$

e então (188) fica

$$i\hbar\Psi(x)\Phi'(t) = -\frac{\hbar^2}{2m}\Psi''(x)\Phi(t) + U(x)\Psi(x)\Phi(t)$$

Dividindo ambos os membros por $\Psi(x)\Phi(t)$, obtemos

$$i\hbar\frac{\Phi'(t)}{\Phi(t)} = -\frac{\hbar^2}{2m}\frac{\Psi''(x)}{\Psi(x)} + U(x) \tag{202}$$

Denotando $[y]$ como a grandeza de $\Phi(t)$, vemos que a grandeza da constante de separação

$$\left[i\hbar\frac{\Phi'(t)}{\Phi(t)}\right] = [J \cdot s] \cdot \frac{[y/s]}{[y]} = [J]$$

é a energia. Denotando a constante de separação de E, temos então duas equações diferenciais ordinárias:

$$i\hbar\Phi'(t) = E \cdot \Phi(t)$$

$$\Phi'(t) = -\frac{iE}{\hbar} \cdot \Phi(t) \tag{203}$$

e

$$-\frac{\hbar^2}{2m}\Psi''(x) + U(x)\Psi(x) = E\Psi(x)$$

$$-\frac{\hbar^2}{2m}\Psi''(x) + [U(x) - E]\Psi(x) = 0 \tag{204}$$

A solução da primeira equação diferencial é

$$\Phi(t) = Ce^{-\frac{iE}{\hbar}t}$$

Já na segunda equação diferencial não podemos ir mais adiante sem especificar o potencial $U(x)$. Ela é a equação de Schrödinger independente do tempo. Entretanto ainda podemos extrair características interessantes dessa solução geral $\psi(x,t) = \Psi(x) \cdot \Phi(t)$ com potencial independente do tempo. Supondo que E seja real, temos que

$$\begin{aligned} |\psi(x,t)|^2 &= \psi(x,t)^*\psi(x,t) \\ &= \Psi(x)^*\Psi(x) \cdot \Phi(t)^*\Phi(t) \\ &= \Psi(x)^*\Psi(x) \cdot Ce^{-\frac{iE}{\hbar}t}C^*e^{\frac{iE}{\hbar}t} \\ &= \Psi(x)^*\Psi(x) \\ &= |\Psi(x)|^2 \end{aligned} \tag{205}$$

e portanto a densidade de probabilidade é invariante em relação ao tempo. Por causa disso, essa função de onda é chamada de estacionária. Daqui em diante deve-se subtender que as funções de onda $\psi(x,t)$ abordadas em toda esta seção são estacionárias.

Em geral, a constante de separação E pode ser um número complexo. Mas se a solução $\psi(x,t)$ for normalizável, então o próximo teorema garante que E é um número real.

Teorema 3.2. *Se $\psi(x,t)$ é uma solução normalizável, então a constante de separação E definida pelo valor da equação* (202) *pertence aos reais.*

Demonstração. Suponha que E seja um complexo na forma

$$E = E_r + iE_i \quad \text{onde} \quad E_r, E_i \in \mathbb{R}$$

Temos que provar que $E_i = 0$ nas condições enunciadas. Então

$$Ce^{-\frac{itE}{\hbar}} = C\exp\left(-\frac{itE_r + it(iE_i)}{\hbar}\right) = C\exp\left(-\frac{iE_r}{\hbar}t\right)\exp\left(\frac{E_i}{\hbar}t\right) \tag{206}$$

e seu conjugado é

$$\left(C\exp\left(-\frac{iE}{\hbar}t\right)\right)^* = C^*\exp\left(\frac{iE_r}{\hbar}t\right)\exp\left(\frac{E_i}{\hbar}t\right) \tag{207}$$

E agora substituindo (206) e (207) em (205) obtemos

$$\begin{aligned} |\psi(x,t)|^2 &= \Psi(x)^*\Psi(x) \cdot C\exp\left(-\frac{iE_r}{\hbar}t\right)C^*\exp\left(\frac{iE_r}{\hbar}t\right)\exp\left(\frac{E_i}{\hbar}t\right)\exp\left(\frac{E_i}{\hbar}t\right) \\ &= \Psi(x)^*\Psi(x) \cdot \exp\left(\frac{2E_i}{\hbar}t\right) \end{aligned}$$

Mas por hipótese $\psi(x,t)$ é normalizável e portanto $|\psi(x,t)|^2$ não pode depender de t. Logo $E_i = 0$, conforme queríamos demonstrar. ■

Teorema 3.3. *Qualquer solução normalizável* $\Psi(x)$ *de* (204) *pode ser escrita por uma combinação linear de outras soluções* $\Psi_i(x)$ *tais que são reais para todo o* x. *Em outras palavras, sempre existirá uma solução real de* (204) *em todo o domínio.*

Demonstração. Pelo teorema (3.2), a constante E é real e consequentemente todos os coeficientes da equação diferencial independente do tempo (204) são reais. Isso significa que podemos aplicar o conjugado em ambos os lados e notar que o conjugado $\Psi(x)^*$ também satisfaz a equação, ou seja, $\Psi(x)^*$ também é solução.

Agora suponha que $\Psi(x)$ pode ser escrito na forma

$$\Psi(x) = a(x) + b(x)i \quad \text{com} \quad a(x), b(x) \in \mathbb{R}$$

Então

$$\Psi(x) - \Psi(x)^* = 2b(x) \in \mathbb{R} \quad \text{e} \quad \Psi(x) + \Psi(x)^* = 2a(x) \in \mathbb{R} \tag{208}$$

Ou seja, podemos escrever

$$\Psi(x) = \frac{1}{2}(\Psi(x) + \Psi(x)^*) + \frac{1}{2}(\Psi(x) - \Psi(x)^*)i$$

Como a EDO (3.2) é linear e homogênea, pelo princípio da superposição as combinações lineares $\Psi(x) - \Psi(x)^*$ e $\Psi(x) + \Psi(x)^*$ também são soluções.

Para obtermos uma solução real, basta usar as equações (208) para qualquer solução normalizável. ■

Teorema 3.4. *Se* $\Psi(x)$ *é uma solução normalizável de* (204)*, então* E *não pode ser menor do que* $\min(U(x))$*.*

Demonstração. Vamos supor por simplicidade que $\Psi(x)$ seja uma função normalizável real. Se não for real, podemos usar o teorema (3.3) para encontrar uma solução real cujo vínculo é dado por uma das equações de (208).

Isolando a derivada de segunda ordem em (204), obtemos

$$\begin{aligned}\Psi''(x) &= -\frac{2m}{\hbar^2}[E - U(x)]\Psi(x)\\ &= \frac{2m}{\hbar^2}[U(x) - E]\Psi(x)\end{aligned}$$

Agora suponha que $E < \min(U(x))$. Então o termo $\frac{2m}{\hbar^2}[U(x) - E]$ é positivo $\forall x \in \mathbb{R}$. Isso significa que ambas as funções $\Psi(x)$ e $\Psi''(x)$ sempre possuem o mesmo sinal. Se $\Psi(x)$ for positivo, então a sua segunda derivada também é. Nesse caso, se $\Psi'(x)$ for positivo, então $\Psi(x)$ tenderá a ∞ quando $x \to \infty$. Se $\Psi'(x)$ for negativo, na melhor das hipóteses $\Psi(x)$ cruzará o eixo x e $\Psi(x)$ se

torna negativo, entretanto $\Psi''(x)$ também será negativo, no que implica $\Psi(x)$ tendendo a $-\infty$ quando $x \to \infty$. De todo modo, $\|\Psi(x)\|^2$ não tenderá a zero quando x tende a $\pm\infty$, o que significa que a integral imprópria

$$\int_{-\infty}^{\infty} \|\Psi(x)\|^2 \, \mathrm{d}x$$

divergirá e não será 1, o que contradiz a hipótese da função ser normalizável. Em outras palavras, o valor mínimo do potencial $U(x)$ é sempre menor do que a energia E do sistema. ■

3.2.4 O problema do poço de potencial infinito

Neste problema o potencial $U(x)$ é definido por

$$U(x) = \begin{cases} 0 & \text{se } 0 \leq x \leq a \\ +\infty & \text{caso contrário} \end{cases}$$

Aqui $+\infty$ deve ser interpretado como um número que tende ao infinito, e não o próprio infinito. Em outras palavras, qualquer que seja o valor finito de $\Psi(x)$, ainda será desprezível comparado ao valor do potencial. Dessa forma, a equação diferencial

$$-\frac{\hbar^2}{2m}\Psi''(x) + [U(x) - E]\Psi(x) = 0 \tag{209}$$

força a condição $\Psi(x) = 0$ para todo x fora do intervalo $[0, a]$, pois a função exata e identicamente nula leva qualquer limite a zero, respeitando (209).

Podemos invocar o teorema (3.4) para descartar $E < 0$. E então, da equação independente do tempo (209) no intervalo $[0, a]$ temos

$$-\frac{\hbar^2}{2m}\Psi''(x) - E\Psi(x) = 0$$

$$\Psi''(x) = \frac{-2mE}{\hbar^2}\Psi(x) = -k^2\Psi(x) \quad \text{onde} \quad k = \frac{\sqrt{2mE}}{\hbar} \in \mathbb{R} \tag{210}$$

Essa é a equação diferencial do oscilador harmônico simples. No caso $E = 0$, temos a solução

$$\Psi(x) = Ax + B$$

e se $E > 0$, temos

$$\Psi(x) = A\cos(kx) + B\sin(kx)$$

Queremos que $\psi(x, t)$ seja contínua, e consequentemente $\Psi(x)$ também precisa ser. Isso automaticamente condiciona os valores de contorno $\Psi(0) = 0$ e $\Psi(a) = 0$. A equação com $E = 0$ nos retorna a função nula, uma vez que é a única função do tipo $Ax + B$ que possui mais de uma raiz. Então nos resta a equação para $E > 0$:

$$\Psi(0) = 0 \implies 0 = A\cos(k0) + B\sin(k0) = A$$

e então ficamos com $\Psi(x) = B\sin(kx)$. Aplicando a segunda condição de contorno:

$$\Psi(a) = B\sin(ka) = 0$$

Não queremos a função nula, então $B \neq 0$ e $\sin(ka) = 0$. Isso significa que

$$ka = n\pi \quad \text{onde} \quad n \in \mathbb{Z}^*$$

Uma vez que a função seno é ímpar, então podemos descartar os valores de n negativos sem prejudicar a generalidade da solução. Isolando k, obtemos

$$k = \frac{n\pi}{a}$$

ou então

$$\frac{\sqrt{2mE_n}}{\hbar} = \frac{n\pi}{a}$$

que leva à expressão

$$E_n = \frac{1}{2m} \cdot \frac{n^2\pi^2\hbar^2}{a^2} \tag{211}$$

Isto é, a energia é quantizada, não podendo assumir valores arbitrários ou contínuos como na mecânica clássica. Esse fato adveio das condições de contorno impostas, condições estas que prezam pela continuidade da função de onda $\psi(x,t)$. Essa quantização também foi vista, conforme mencionamos em (3.1), no caso dos osciladores térmicos da radiação do corpo negro.

A solução geral é dada, portanto, pela função

$$\Psi(x) = \begin{cases} \displaystyle\sum_{n=1}^{\infty} B_n \sin\left(\frac{n\pi x}{a}\right) & \text{se } x \in [0,a] \\ 0 & \text{caso contrário} \end{cases} \tag{212}$$

Mas ainda resta a normalização. Calculando $|\Psi(x)|^2$:

$$\|\Psi(x)\|^2 = \left(\sum_{n=1}^{\infty} B_n \sin\left(\frac{n\pi x}{a}\right)\right) \cdot \left(\sum_{n=1}^{\infty} B_n \sin\left(\frac{n\pi x}{a}\right)\right)$$

Do produto de Cauchy $\left(\displaystyle\sum_{n=1}^{\infty} a_n\right) \cdot \left(\displaystyle\sum_{n=1}^{\infty} b_n\right) = \displaystyle\sum_{n=1}^{\infty}\sum_{j=1}^{n} a_k b_{n-j}$, temos

$$\|\Psi(x)\|^2 = \sum_{n=1}^{\infty}\sum_{j=1}^{n} B_j B_{n-j} \sin\left(\frac{j\pi x}{a}\right) \sin\left(\frac{(n-j)\pi x}{a}\right)$$

A hipótese que fizemos no começo do cálculo é de que $\Psi(x) = 0$ para qualquer $x \notin [0,a]$, então podemos encurtar o intervalo de integração para $[-a,a]$ sem prejudicar a norma. O subintervalo $[-a,0)$ (cuja densidade é zero) vai estar incluso por questões de conveniência.

$$\int_{-\infty}^{\infty} |\Psi(x)|^2 \, \mathrm{d}x = \int_{-a}^{a} \sum_{n=1}^{\infty} \sum_{j=1}^{n} B_j B_{n-j} \sin\left(\frac{j\pi x}{a}\right) \sin\left(\frac{(n-j)\pi x}{a}\right) \mathrm{d}x$$

$$= \sum_{n=1}^{\infty} \sum_{j=1}^{n} B_j B_{n-j} \int_{-a}^{a} \sin\left(\frac{j\pi x}{a}\right) \sin\left(\frac{(n-j)\pi x}{a}\right) \mathrm{d}x$$

Conforme vimos no começo da seção (2.5), podemos usar a relação de ortogonalidade para mostrar que

$$\int_{-a}^{a} \sin\left(\frac{j\pi x}{a}\right) \sin\left(\frac{(n-j)\pi x}{a}\right) \mathrm{d}x = 0 \quad \text{se} \quad j \neq n - j$$

ou seja, os únicos termos não nulos das somatórias são os que $j = n - j$. Em outras palavras, só sobram termos que tenham n par e ao mesmo tempo $2j = n$. E então

$$\int_{-\infty}^{\infty} |\Psi(x)|^2 \, \mathrm{d}x = \sum_{j=1}^{\infty} B_j B_j \int_{-a}^{a} \sin\left(\frac{j\pi x}{a}\right) \sin\left(\frac{j\pi x}{a}\right) \mathrm{d}x$$

$$= \sum_{j=1}^{\infty} B_j^2 \int_{-a}^{a} \sin^2\left(\frac{j\pi x}{a}\right) \mathrm{d}x$$

Agora vamos diminuir o intervalo de integração de $[-a, a]$ para $[0, a]$.

$$\int_{-\infty}^{\infty} |\Psi(x)|^2 \, \mathrm{d}x = \sum_{j=1}^{\infty} B_j^2 \int_{0}^{a} \sin^2\left(\frac{j\pi x}{a}\right) \mathrm{d}x$$

$$= \sum_{j=1}^{\infty} B_j^2 \int_{0}^{a} \frac{1}{2} - \frac{1}{2} \cdot \cos\left(\frac{2j\pi x}{a}\right) \mathrm{d}x$$

$$= \sum_{j=1}^{\infty} B_j^2 \left[\frac{x}{2} - \frac{a \sin\left(\frac{2j\pi x}{a}\right)}{2 \cdot 2j\pi} \right] \Bigg|_{0}^{a}$$

$$= \sum_{j=1}^{\infty} B_j^2 \left[\frac{a}{2} - \frac{a \sin(2j\pi)}{2 \cdot 2j\pi} \right]$$

$$= \frac{a}{2} \sum_{j=1}^{\infty} B_j^2$$

A condição de normalização (190) implica

$$\frac{a}{2} \sum_{j=1}^{\infty} B_j^2 = 1$$

$$\sum_{j=1}^{\infty} B_j^2 = \frac{2}{a} \tag{213}$$

ou seja, as constantes B_j devem ser escolhidas de modo a satisfazer (213).

Essa é a solução que engloba todas as energias possíveis pelo sistema. No caso particular em que queremos uma função de onda para cada energia, temos

$$\Psi_n(x) = B_n \sin\Big(\frac{n\pi x}{a}\Big) \quad \text{se} \quad x \in [0, a]$$

E a equação (213) simplifica para

$$B_n^2 = \frac{2}{a}$$

$$B_n = \sqrt{\frac{2}{a}}$$

Escolhemos a raiz positiva por questões de conveniência. O sinal de B_n não interessa muito, visto que é usado B_n^2 para o cálculo da normalização. E então a solução em $[0, a]$ fica

$$\Psi_n(x) = \sqrt{\frac{2}{a}} \sin\Big(\frac{n\pi x}{a}\Big)$$

Cada solução $\Psi_n(x)$ está associado à energia dada pela equação (214). A solução com energia associada a $n = 1$ é chamada de estado fundamental. As soluções com $n > 1$ são chamadas de estados excitados.

$$E_n = \frac{1}{2m} \cdot \frac{n^2 \pi^2 \hbar^2}{a^2} \tag{214}$$

Voltando à solução global (212), podemos generalizar B_n para os complexos e reescrever-la como

$$\Psi(x) = \sqrt{\frac{2}{a}} \sum_{n=1}^{\infty} C_n \sin\Big(\frac{n\pi x}{a}\Big) \quad \text{onde} \quad B_n = \sqrt{\frac{2}{a}} C_n \tag{215}$$

Agora dada uma outra solução particular $\Psi_m(x)$ em $[a, b]$

$$\Psi_m(x) = \sqrt{\frac{2}{a}} \sin\Big(\frac{m\pi x}{a}\Big)$$

vamos multiplicar ambos os lados de (215) pelo conjugado de $\Psi_m(x)$ e integrar:

$$\begin{aligned} \int_0^a \Psi(x)\Psi_m(x)^* \,\mathrm{d}x &= \int_0^a \frac{2}{a} \sum_{n=1}^{\infty} C_n \sin\Big(\frac{n\pi x}{a}\Big) \sin\Big(\frac{m\pi x}{a}\Big) \,\mathrm{d}x \\ &= \frac{2}{a} C_m \int_0^a \sin^2\left(\frac{m\pi x}{a}\right) \mathrm{d}x \\ &= \frac{2}{a} \cdot \frac{a}{2} C_m \\ &= C_m \end{aligned} \tag{216}$$

Agora note que a integral de $\Psi(x)\Psi_m(x)^*$ é o produto interno definido entre funções complexas:

$$\langle \Psi_m(x), \Psi(x) \rangle = \int_0^a \Psi(x)\Psi_m(x)^* \, \mathrm{d}x \tag{217}$$

Consequentemente as constantes C_n são as coordenadas da projeção da solução geral $\Psi(x)$ no espaço gerado pelas funções $\Psi_m(x)$. Note que essas soluções são linearmente independentes e assim formam uma base (infinita) β. Se observarmos bem a equação de Schrödinger independente do tempo (210) percebemos que cada solução $\Psi_n(x)$ é um autovetor do operador da derivada de segunda ordem com autovalor

$$\lambda_n = \frac{-2nE_n}{\hbar^2}$$

Além disso, se fizermos o produto interno entre dois elementos da base β vamos obter

$$\begin{aligned}\langle \Psi_m(x), \Psi_n(x) \rangle &= \int_0^a \Psi_n(x)\Psi_m(x)^* \, \mathrm{d}x \\ &= \frac{2}{a}\int_0^a \sin\Big(\frac{n\pi x}{a}\Big)\sin\Big(\frac{m\pi x}{a}\Big)\, \mathrm{d}x\end{aligned}$$

Isto é, se $n \neq m$, o produto interno vale zero. E se $n = m$, a integral vale $\frac{a}{2}$ e então o produto interno é 1. Em outras palavras, a base β é ortonormal, conforme já vimos no exemplo (1.3) e na seção (2.5):

$$\langle \Psi_m(x), \Psi_n(x) \rangle = \delta_{nm}$$

Por fim, a solução geral no intervalo $[0, a]$ é

$$\Psi(x) = \sqrt{\frac{2}{a}} \sum_{n=1}^{\infty} C_n \sin\Big(\frac{n\pi x}{a}\Big)$$

que é a Série de Fourier de alguma função ímpar arbitrária. Por sorte, as constantes C_n podem ser calculadas pelas equações da projeção de f no espaço gerador da Série de Fourier:

$$\sqrt{\frac{2}{a}} C_n = \frac{1}{a}\int_{-a}^{a} f(x)\sin\Big(\frac{n\pi x}{a}\Big)\, \mathrm{d}x$$

$$C_n = \sqrt{\frac{1}{2a}}\int_{-a}^{a} f(x)\sin\Big(\frac{n\pi x}{a}\Big)\, \mathrm{d}x$$

onde f é uma função ímpar de livre escolha desde que obedeça aos critérios da convergência de uniforme de (2.35) ou (2.36). Podemos simplificar mais ainda,

uma vez que $f(x)\sin\left(\frac{n\pi x}{a}\right)$ é par:

$$\begin{aligned} C_n &= \frac{2}{\sqrt{2}}\frac{1}{\sqrt{a}}\int_0^a f(x)\sin\left(\frac{n\pi x}{a}\right)\mathrm{d}x \\ &= \sqrt{\frac{2}{a}}\,\langle \Psi_n(x), f(x)\rangle \end{aligned}$$

Combinando essa expressão com (217), obtemos

$$C_n = \langle \Psi_n(x), \Psi(x)\rangle = \sqrt{\frac{2}{a}}\,\langle \Psi_n(x), f(x)\rangle$$

3.3 Introdução ao oscilador harmônico quântico

Seja $U(x)$ uma função analítica de um potencial que possui um mínimo local em $x = x_0$. Estamos interessados em saber como a função de onda $\psi(x,t)$ se comporta nas vizinhanças de x_0. Essa análise é útil, por exemplo, em problemas envolvendo interações moleculares, usando o potencial de Lennard-Jones.

Podemos fazer uma aproximação polinomial de Taylor centrada em x_0 e ignorar os termos de ordem mais alta:

$$U(x) \approx U(x_0) + U'(x_0)\cdot(x - x_0) + \frac{1}{2}U''(x_0)\cdot(x - x_0)^2$$

Dado que x_0 é um mínimo local, então $U'(x_0) = 0$. Além disso, para a física em geral apenas interessa a diferença de potencial, que é invariante a soma ou subtração de uma constante. Então podemos ignorar o termo $U(x_0)$ e concluir que

$$U(x) \approx \frac{1}{2}U''(x_0)\cdot(x - x_0)^2$$

ou seja, o potencial nas vizinhanças de x_0 possui a mesma forma que o potencial do oscilador harmônico simples da mecânica clássica $\frac{1}{2}k(x - x_0)^2$ centrado em x_0 e com a constante da mola $k = U''(x_0)$.

Sabemos que no caso clássico a frequência angular ω é dado por

$$\omega = \sqrt{\frac{k}{m}}$$

e então podemos isolar k para obter

$$k = m\omega^2$$

E portanto podemos reescrever o potencial como

$$U(x) = \frac{1}{2}m\omega^2(x - x_0)^2$$

Colocando esse potencial em (204), obtemos

$$-\frac{\hbar^2}{2m}\Psi''(x - x_0) + \left[\frac{1}{2}m\omega^2(x - x_0)^2 - E\right]\Psi(x - x_0) = 0$$

Multiplicando ambos os lados por $-\dfrac{2m}{\hbar^2}$, temos

$$\Psi''(x-x_0)+\left[\frac{2mE}{\hbar^2}-\frac{m^2}{\hbar^2}\omega^2(x-x_0)^2\right]\Psi(x-x_0)=0$$

Note que o primeiro termo da expressão em colchetes é o quadrado de $k=\dfrac{\sqrt{2mE}}{\hbar}$, também definido na equação diferencial do poço infinito (210).

Agora vamos definir uma outra variável ξ:

$$\xi=\sqrt{\frac{m\omega}{\hbar}}(x-x_0)$$

$$\frac{d\xi}{dx}=\sqrt{\frac{m\omega}{\hbar}}$$

$$\begin{aligned}\frac{d^2\Psi(x)}{dx^2}&=\frac{d}{dx}\left(\frac{d\Psi(\xi)}{d\xi}\cdot\frac{d\xi}{dx}\right)\\&=\frac{d^2\Psi(\xi)}{d\xi^2}\cdot\frac{d\xi}{dx}\cdot\frac{d\xi}{dx}+\frac{d\Psi(\xi)}{d\xi}\cdot\frac{d^2\xi}{dx^2}\\&=\frac{d^2\Psi(\xi)}{d\xi^2}\left(\frac{d\xi}{dx}\right)^2\\&=\frac{m\omega}{\hbar}\frac{d^2\Psi(\xi)}{d\xi^2}\end{aligned}$$

e então a equação diferencial fica

$$\frac{m\omega}{\hbar}\Psi''(\xi)+\left[\frac{2mE}{\hbar^2}-\frac{m\omega}{\hbar}\xi^2\right]\Psi(\xi)=0$$

Dividindo ambos os lados por $\dfrac{m\omega}{\hbar}$ resulta

$$\Psi''(\xi)+\left[\frac{2E}{\hbar\omega}-\xi^2\right]\Psi(\xi)=0$$

$$\Psi''(\xi)+\left[K-\xi^2\right]\Psi(\xi)=0\quad\text{onde}\quad K=\frac{2E}{\hbar\omega}\tag{218}$$

Quando ξ^2 é muito grande comparado à constante K, podemos desprezar K e ficamos com

$$\Psi''(x)=\xi^2\Psi(\xi)\tag{219}$$

No contexto em que $\xi^2>>K$, $\Psi(\xi)$ não pode ter um comportamento polinomial, uma vez que a derivada diminui o grau e não o contrário. Uma exponencial do tipo $e^{\lambda\xi}$ também não pode ser, já que ξ^2 não é uma constante. E para a exponencial $e^{\lambda\xi^2}$ temos

$$\frac{d}{d\xi}e^{\lambda\xi^2} = 2\lambda\xi e^{\lambda\xi^2}$$

$$\frac{d^2}{d\xi^2}e^{\lambda\xi^2} = \left(2\lambda + 4\lambda^2\xi^2\right)e^{\lambda\xi^2}$$

Mas como ξ^2 é muito grande, podemos desprezar o termo 2λ e concluir que

$$\frac{d^2}{d\xi^2}e^{\lambda\xi^2} \approx 4\lambda^2\xi^2 e^{\lambda\xi^2}$$

Para se assemelhar à equação diferencial (219), fazemos $4\lambda^2 = 1$, ou então

$$\lambda = \pm\frac{1}{2} \qquad \lambda_0 = -\frac{1}{2} \qquad \lambda_1 = \frac{1}{2}$$

A solução da equação diferencial (218) que estamos procurando precisa, então, convergir para a solução

$$\Psi(\xi) \approx Ae^{\lambda_0\xi^2} + Be^{\lambda_1\xi^2} \tag{220}$$

quando $\xi^2 >> K$. Para que a solução $\Psi(\xi)$ seja quadrado integrável de $-\infty$ a ∞, é condição necessária

$$\lim_{\xi\to\infty}\Psi(\xi) = 0 \quad \text{e} \quad \lim_{\xi\to-\infty}\Psi(\xi) = 0$$

Quando ξ^2 tende a infinito em (220), o termo $Be^{\lambda_1\xi^2}$ diverge, enquanto que $Ae^{\lambda_0\xi^2}$ converge para zero. E então B precisa ser zero. Em outras palavras, $\Psi(\xi)$ tende a função $Ae^{\lambda_0\xi^2}$ quando ξ^2 é grande.

Isso nos fornece pistas de que $\Psi(\xi)$ pode ser reescrita na forma

$$\Psi(\xi) = f(\xi)e^{\lambda_0\xi^2} \tag{221}$$

Derivando uma e duas vezes:

$$\Psi'(\xi) = 2\lambda_0\xi f(\xi)e^{\lambda_0\xi^2} + f'(\xi)e^{\lambda_0\xi^2}$$

$$\begin{aligned}\Psi''(\xi) &= 2\lambda_0 f(\xi)e^{\lambda_0\xi^2} + 4\lambda_0^2\xi^2 f(\xi)e^{\lambda_0\xi^2} + 2\cdot 2\lambda_0\xi f'(\xi)e^{\lambda_0\xi^2} + f''(\xi)e^{\lambda_0\xi^2} \\ &= \left[(4\lambda_0^2\xi^2 + 2\lambda_0)f(\xi) + 2\cdot 2\lambda_0\xi f'(\xi) + f''(\xi)\right]e^{\lambda_0\xi^2}\end{aligned}$$

Substituindo $\Psi''(\xi)$ e $\Psi(\xi)$ em (218), obtemos

$$\left[(4\lambda_0^2\xi^2 + 2\lambda_0)f(\xi) + 2\cdot 2\lambda_0\xi f'(\xi) + f''(\xi)\right]e^{\lambda_0\xi^2} = [\xi^2 - K]f(x)e^{\lambda_0\xi^2}$$

A exponencial nunca é zero, então podemos retirá-la da equação dividindo por ela:

$$(4\lambda_0^2\xi^2 + 2\lambda_0)f(\xi) + 2\cdot 2\lambda_0\xi f'(\xi) + f''(\xi) = [\xi - K]f(\xi)$$

$$\begin{aligned}
0 &= f''(\xi) + 4\lambda_0 \xi f'(\xi) + [4\lambda_0^2 \xi^2 - \xi^2 + K + 2\lambda_0] f(\xi) \\
&= f''(\xi) + 4\lambda_0 \xi f'(\xi) + \left[4\frac{1}{4}\xi^2 - \xi^2 + K + 2\lambda_0\right] f(\xi) \\
&= f''(\xi) + 4\lambda_0 \xi f'(\xi) + (K + 2\lambda_0) f(\xi) \qquad (222)
\end{aligned}$$

Sabemos que

$$4\lambda_0 = -4 \cdot \frac{1}{2} = -2 \quad \text{e} \quad 2\lambda_0 = -1$$

e portanto a equação diferencial se torna

$$f''(\xi) - 2\xi f'(\xi) + (K - 1) f(\xi) = 0 \qquad (223)$$

Como os coeficientes -2ξ e $(K-1)$ são analíticos em $\mathbb{R}$, então essa equação diferencial admite duas soluções analíticas linearmente independentes (demonstração na seção 6.17 de [5]). Logo

$$f(\xi) = \sum_{n=0}^{\infty} a_n \xi^n$$

$$f'(\xi) = \sum_{n=0}^{\infty} n a_n \xi^{n-1}$$

$$f''(\xi) = \sum_{n=0}^{\infty} n(n-1) a_n \xi^{n-2} = \sum_{n=2}^{\infty} n(n-1) a_n \xi^{n-2} = \sum_{n=0}^{\infty} (n+2)(n+1) a_{n+2} \xi^n$$

Substituindo em (222), temos

$$\begin{aligned}
0 &= \sum_{n=0}^{\infty} (n+2)(n+1) a_{n+2} \xi^n - 2\xi \sum_{n=0}^{\infty} n a_n \xi^{n-1} + (K-1) \sum_{n=0}^{\infty} a_n \xi^n \\
&= \sum_{n=0}^{\infty} [(n+2)(n+1) a_{n+2} + (-2n + K - 1) a_n] \cdot \xi^n
\end{aligned}$$

Como os monômios ξ^n são linearmente independentes, então o termo entre colchetes tem que ser zero para todo $n \geq 0$:

$$(n+2)(n+1) a_{n+2} + (-2n + K - 1) a_n = 0$$

$$a_{n+2} = \frac{(2n + 1 - K) a_n}{(n+2)(n+1)} \qquad (224)$$

A equação (224) é a relação de recorrência que permite determinar todos os coeficientes a_n. Rapidamente podemos notar que a_0 e a_1 são arbitrários, e os termos seguintes dependerão da escolha de a_0 e a_1. Calculando os primeiros coeficientes:

$$a_2 = a_{0+2} = \frac{(1 - K) a_0}{2!}$$

$$a_3 = a_{1+2} = \frac{(3-K)a_1}{3!}$$
$$a_4 = a_{2+2} = \frac{(5-K)a_2}{4 \cdot 3} = \frac{(5-K)(1-K)a_0}{4!}$$
$$a_5 = a_{3+2} = \frac{(7-K)a_3}{5 \cdot 4} = \frac{(7-K)(3-K)a_1}{5!}$$
$$a_6 = \frac{(9-K)a_4}{6 \cdot 5} = \frac{(9-K)(5-K)(1-K)a_0}{6!}$$
$$a_7 = \frac{(11-K)a_5}{7 \cdot 6} = \frac{(11-K)(7-K)(3-K)a_1}{7!}$$

Esses coeficientes são suficientes para mostrar um padrão em que a cada dois coeficientes, um novo produto surge, que é o anterior somado por 4. Os coeficientes pares dependem de a_0 e os ímpares de a_1. Além disso, O coeficiente a_{2n} tem n produtos, e a_{2n+1} tem também n produtos. Portanto é plausível tentar

$$a_{2n} = \frac{a_0}{(2n)!} \cdot \prod_{j=0}^{n-1} ((4j+1) - K) \quad \text{se} \quad n \geq 1 \tag{225}$$

e

$$a_{2n+1} = \frac{a_1}{(2n+1)!} \cdot \prod_{j=0}^{n-1} ((4j+3) - K) \quad \text{se} \quad n \geq 1 \tag{226}$$

Para provar essas expressões, é preciso usar o princípio da indução. Para os coeficientes pares:

$$\begin{aligned}
a_{2(n+1)} &= \frac{a_0}{(2(n+1))!} \cdot \prod_{j=0}^{n} ((4j+1) - K) \\
&= \frac{((4n+1) - K)a_0}{(2n+2)!} \cdot \prod_{j=0}^{n-1} ((4j+1) - K) \\
&= \frac{((4n+1) - K)}{(2n+1) \cdot (2n+2)} \cdot \frac{a_0}{(2n)!} \cdot \prod_{j=0}^{n-1} ((4j+1) - K) \\
&= \frac{((4n+1) - K)a_{2n}}{(2n+1) \cdot (2n+2)}
\end{aligned}$$

que é a relação de recorrência (224) se substituirmos n por $2n$, e portanto está provado por indução.

E para os coeficientes ímpares:

$$\begin{aligned}
a_{2(n+1)+1} &= \frac{a_1}{(2(n+1)+1)!} \cdot \prod_{j=0}^{n}((4j+3)-K) \\
&= \frac{((4n+3)-K)a_1}{(2n+3)!} \cdot \prod_{j=0}^{n-1}((4j+3)-K) \\
&= \frac{((4n+3)-K)}{(2n+2)\cdot(2n+3)} \cdot \frac{a_1}{(2n+1)!} \cdot \prod_{j=0}^{n-1}((4j+3)-K) \\
&= \frac{((4n+3)-K)a_{2n+1}}{(2n+2)\cdot(2n+3)}
\end{aligned}$$

e retornamos à relação (224) substituindo n por $2n+1$.

Portanto a solução f é dada por

$$f(\xi) = \sum_{n=0}^{\infty} a_{2n}\xi^{2n} + \sum_{n=0}^{\infty} a_{2n+1}\xi^{2n+1}$$

com a_{2n} e a_{2n+1} definidos respectivamente pelas equações (225) e (226).

Em uma análise mais grosseira, quando n é um número muito grande, podemos arredondar a relação de recorrência dos coeficientes pares para

$$a_{2n+2} \approx \frac{2(2n)a_n}{2n(2n+2)} = \frac{1}{n+1}a_{2n} \tag{227}$$

de forma que

$$a_2 \approx a_0 \qquad a_4 \approx \frac{1}{2}a_2 = \frac{a_0}{2}$$
$$a_6 \approx \frac{a_4}{3} = \frac{a_0}{2\cdot 3} \qquad a_8 \approx \frac{a_0}{4!}$$

Esse padrão nos sugere que

$$a_{2n} \approx \frac{a_0}{n!} \tag{228}$$

Substituindo na expressão (227), obtemos

$$a_{2n+2} \approx \frac{1}{n+1}\frac{a_0}{n!} = \frac{a_0}{(n+1)!}$$

que é (228) trocando n por $n+1$ e consequentemente a relação está provada.

Portanto, para n muito grande,

$$f(\xi) = \sum_{n=0}^{\infty} a_{2n}\xi^{2n} \approx \sum_{n=0}^{\infty} \frac{a_0}{n!}\xi^{2n} \approx \sum_{n=0}^{\infty} \frac{a_0}{n!}(\xi^2)^n \approx a_0 e^{\xi^2}$$

E então

$$\Psi(\xi) = f(\xi)e^{-\xi^2/2} \approx Ce^{\xi^2}e^{-\xi^2/2} \approx Ce^{\xi^2/2}$$

Isso significa que a função $\Psi(\xi)$ diverge quando ξ^2 tende a infinito, e consequentemente não é normalizável se $a_0 \neq 0$. Para os coeficientes ímpares, o resultado será semelhante.

Para que $\Psi(\xi)$ seja normalizável, é preciso que as séries infinitas sejam na verdade finitas, ou seja, polinômios. Conforme já sabemos, as exponenciais crescem (ou decrescem) mais rapidamente do que qualquer polinômio, e com isso $\Psi(\xi)$ tende a $e^{\lambda_0 \xi^2}$ quando ξ^2 é muito grande.

Vendo a relação de recorrência

$$a_{n+2} = \frac{(2n+1-K)a_n}{(n+2)(n+1)} \tag{229}$$

a série é truncada quando ou $a_n = 0$ ou $2n+1-K = 0$ para um certo n. Por hipótese, os coeficientes anteriores não são zero e então $2n+1-K = 0$. Ou seja,

$$K = 2n+1$$

$$\frac{2E}{\hbar\omega} = 2n+1$$

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right)$$

Ou seja, para que $\psi(x,t)$ seja normalizável e possa caracterizar uma partícula, é preciso que a energia (mais uma vez) seja quantizada.

Quando o n de corte é zero, temos $E_0 = \frac{1}{2}\hbar\omega$ e $(2 \cdot 0 + 1 - K) = 0$, no que acarreta

$$a_2 = 0$$

e idem para os outros coeficientes pares (exceto a_0). A única maneira da série infinita dos coeficientes ímpares não explodir quando $\xi^2 \to \infty$ é impondo $a_1 = 0$. Então $f(\xi) = a_0$ e

$$\Psi_0(\xi) = a_0 e^{\lambda_0 \xi^2} \qquad \text{onde} \qquad \lambda_0 = -\frac{1}{2}$$

Quando o n de corte é 1, $E_1 = \frac{3}{2}\hbar\omega$ e os próximos coeficientes ímpares, a_3 em diante, serão nulos. E então a série dos coeficientes pares só pode ser finita se $a_0 = 0$. Além disso, a solução para f é $f(\xi) = a_1 \xi$ enquanto que para Ψ é

$$\Psi_1(\xi) = a_1 \xi e^{\lambda_0 \xi^2}$$

Por questão de conveniência (que será esclarecido posteriormente), multiplicaremos essa solução por 2 para que o monômio com a potência mais alta fique com um fator 2^n, resultando

$$\Psi_1(\xi) = 2a_1 \xi e^{\lambda_0 \xi^2}$$

O padrão fica evidente quando fazemos $n = 2$: os termos pares anteriores (a_0 e o próprio a_2) não são necessariamente nulos enquanto a_4 em diante são

zero. Além disso, a_1 também precisa ser zero para a solução ser normalizável. O valor a_0 é independente, mas a_2 é dependente de a_0 por meio da relação de recorrência (229):

$$a_2 = \frac{1}{2}(1 - K_2)a_0$$

onde

$$\begin{aligned} K_2 &= \frac{2E_2}{\hbar\omega} \\ &= \frac{2}{\hbar\omega}\frac{5\hbar\omega}{2} \\ &= 5 \end{aligned}$$

$$a_2 = \frac{1}{2}(1-5)a_0 = -2a_0$$

E então a solução fica

$$\begin{aligned} \Psi_2(\xi) &= (a_0 + a_2\xi^2)e^{\lambda_0\xi^2} \\ &= a_0\left(1 - 2\xi^2\right)e^{\lambda_0\xi^2} \\ &= -\frac{a_0}{2}\left(4\xi^2 - 2\right)e^{\lambda_0\xi^2} \end{aligned}$$

Podemos simplificar o termo $(q - K_n)$ como

$$\begin{aligned} q - K_n &= q - \frac{2}{m\omega}\cdot\left(n + \frac{1}{2}\right)m\omega \\ &= q - (2n+1) \end{aligned}$$

Para $n = 3, 4$ e 5:

$$\begin{aligned} \Psi_3(\xi) &= (a_1\xi + a_3\xi^3)e^{\lambda_0\xi^2} \\ &= a_1\left(\xi - \frac{2}{3}\xi^3\right)e^{\lambda_0\xi^2} \\ &= -\frac{a_1}{12}(8\xi^3 - 12\xi)e^{\lambda_0\xi^2} \end{aligned}$$

$$\begin{aligned} \Psi_4(\xi) &= (a_0 + a_2x^2 + a_4x^4)e^{\lambda_0\xi^2} \\ &= a_0\left(1 - 4\xi^2 + \frac{4}{3}\xi^4\right)e^{\lambda_0\xi^2} \\ &= \frac{a_0}{12}\left(12 - 48\xi^2 + 16\xi^4\right)e^{\lambda_0\xi^2} \end{aligned}$$

$$\begin{aligned} \Psi_5(\xi) &= (a_1\xi + a_3\xi^3 + a_5\xi^5)e^{\lambda_0\xi^2} \\ &= a_1\left(\xi - \frac{4}{3}\xi^3 + \frac{4}{15}\xi^5\right)e^{\lambda_0\xi^2} \\ &= \frac{a_1}{120}(120\xi - 160\xi^3 + 32\xi^5)e^{\lambda_0\xi^2} \end{aligned}$$

Note que os polinômios entre parênteses de cada solução particular são os polinômios de Hermite do exemplo (1.5) multiplicados por 2^n.

A solução particular da EDO (223) é dada por

$$\Psi_n(\xi) = b_n H_n(\xi) e^{\lambda_0 \xi^2}$$

onde H_n são os polinômios de Hermite cuja potência mais alta tem um fator de 2^n. Quanto a normalização dessas soluções, é usado o fato de que os polinômios de Hermite são ortogonais entre si em relação ao produto interno definido em (1.5)

$$\langle H_n(\xi), H_m(\xi) \rangle = \int_{-\infty}^{\infty} e^{\lambda_0 \xi^2} H_n(\xi) H_m(\xi) \, \mathrm{d}\xi \quad \text{onde} \quad \lambda_0 = -\frac{1}{2}$$

e a fórmula de recorrência para esses polinômios

$$H_{n+1}(\xi) = 2\xi H_n(\xi) - H_n'(\xi)$$

para mostrar que a escolha

$$a_n = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \cdot \frac{1}{\sqrt{2^n n!}}$$

satisfaz a condição de normalização

$$\int_{-\infty}^{\infty} |\Psi_n(x)|^2 \, \mathrm{d}x = 1 \quad \text{onde} \quad x = \sqrt{\frac{\hbar}{m\omega}}\xi + x_0$$

Os detalhes da normalização podem ser vistos na seção 13 da referência [6].

As cinco primeiras soluções normalizadas são mostradas no gráfico da figura (6).

Como a equação diferencial (223) é homogênea e linear, então qualquer combinação linear das soluções particulares também será uma solução. Além disso, os polinômios de Hermite são linearmente independentes, uma vez que são ortogonais entre si. Então podemos englobar todas as soluções particulares em uma só: a solução geral é

$$\Psi(\xi) = \sum_{n=0}^{\infty} c_n H_n e^{-\xi^2/2}$$

No problema do poço de potencial infinito, usamos a Série de Fourier para escrever a condição inicial $\psi(x,0) = \sqrt{\frac{2}{a}} f(x)$ em termos de senos, porém no caso do oscilador harmônico quântico é preciso escrever a condição inicial $f(\xi)$ em termos de $H_n e^{-\xi^2/2}$, escolhendo corretamente as constantes b_n caso isso seja possível.
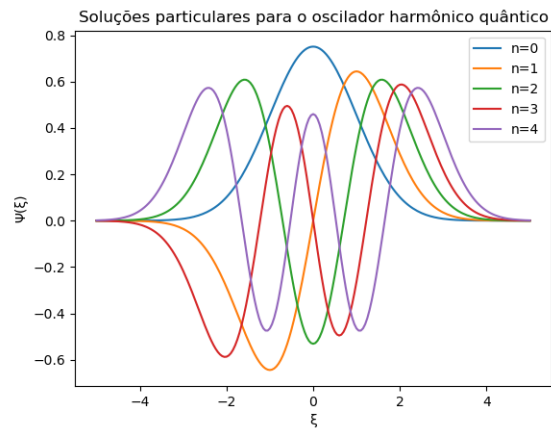
Figura 6: As cinco primeiras soluções normalizadas para o problema do oscilador harmônico quântico, considerando $m = 1$, $\omega = 1$ e $\hbar = 1$.

4 Teoria de Sturm-Liouville

4.1 Distribuição delta de Dirac

O delta de Kronecker δ_{ij}, podendo ser interpretado como um tensor de segunda ordem ou como uma função de i e j, é um objeto que satisfaz a relação

$$\sum_j \delta_{ij} a_j = a_i$$

na qual é útil e foi usada diversas vezes nas seções passadas. Um análogo contínuo do delta de Kronecker é o delta de Dirac, que satisfaz

$$\int_b^c \delta(x,a) f(x)\,\mathrm{d}x = f(a)$$

onde $a \in [b,c]$.

O delta de Dirac é usado, por exemplo, no cálculo do fluxo de um campo gravitacional $\vec{g}$ sobre uma superfície fechada S, produzido por um ponto P localizado pelo vetor $\vec{r_0}$. O campo é dado por

$$\vec{g} = -\frac{Gm}{\|\vec{r} - \vec{r_0}\|^3}(\vec{r} - \vec{r_0})$$

Se essa superfície contém P em seu interior (ou seja, P não está na fronteira de S), então o fluxo é calculado por $-4\pi Gm$, mas caso contrário, o fluxo é zero. Além disso, como $\vec{g}$ é contínuo e diferenciável nas fronteiras de S, pelo teorema de Gauss:

$$\oiint_S \vec{g} \cdot \mathrm{d}\vec{\sigma} = \iiint_V \vec{\nabla} \cdot \vec{g}\,\mathrm{d}V = -4\pi G \iiint_V \rho\,\mathrm{d}V$$

Mas como o fluxo é zero se a superfície S não englobar $\vec{r}_0$, segue que

$$\rho = M\delta(\vec{r}, \vec{r}_0)$$

Se definirmos uma função g tal que $f(x) \neq g(x)$ para todo $x \neq a$ em $[b,c]$ mas $f(a) = g(a)$, então

$$\int_b^c \delta(x,a)(f(x) - g(x))\,\mathrm{d}x = f(a) - g(a) = 0$$

resultado que não importa o quanto f e g diferem entre si nos pontos em que $x \neq a$. Dai podemos inferir que $\delta(x,a) = 0$ quando $x \neq a$ e que $\delta(x,a)$ possui uma singularidade em $x = a$.

No caso especial em que $f(x) = 1 \quad \forall x$, temos

$$\int_b^c \delta(x,a)\,\mathrm{d}x = 1$$

Com essa propriedade, podemos definir o delta de Dirac $\delta(x,a)$ como sendo o limite de qualquer sequência de funções $\{h_n(x)\}$ que satisfaça duas propriedades:

- para todo $x \neq a$, o limite $\lim_{n\to\infty} h_n(x)$ convergirá pontualmente para zero;
- o limite $\lim_{n\to\infty} \int_{-\infty}^{\infty} h_n(x)\,\mathrm{d}x$ convergirá para 1.

Dessa forma, se $c < a < b$, então

$$\begin{aligned}
\lim_{n\to\infty} \int_{c}^{b} f(x)h_n(x)\,\mathrm{d}x &= \lim_{n\to\infty} \int_{a-\epsilon/2}^{a+\epsilon/2} f(x)h_n(x)\,\mathrm{d}x \\
&= \lim_{n\to\infty} \int_{a-\epsilon/2}^{a+\epsilon/2} f(a)h_n(x)\,\mathrm{d}x \\
&= f(a) \lim_{n\to\infty} \int_{a-\epsilon/2}^{a+\epsilon/2} h_n(x)\,\mathrm{d}x \\
&= f(a)
\end{aligned}$$

considerando que $f(x) \approx f(a)$ no intervalo $[a-\epsilon/2, a+\epsilon/2]$ quando $\epsilon \to 0$.

Exemplo 4.1. *Seja $\{h_n(x)\}$ uma sequência de funções, onde*

$$h_n(x) = \begin{cases} \dfrac{n}{2} & \text{se } |x-a| < \dfrac{1}{n} \\ 0 & \text{caso contrário} \end{cases}$$

A sequência $\{h_n(x)\}$ satisfaz

$$\lim_{n\to\infty} h_n(x) = 0 \quad \text{se} \quad x \neq a$$

pois para qualquer $x \neq a$, existe um N tal que para todo $n \geq N$, $|x-a| > \dfrac{1}{n}$ e portanto $h_n(x) = 0$. Além disso,

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} h_n(x)\,\mathrm{d}x =$$

$$\lim_{n\to\infty} \int_{-\infty}^{-\frac{1}{n}} h_n(x)\,\mathrm{d}x + \lim_{n\to\infty} \int_{-\frac{1}{n}}^{\frac{1}{n}} h_n(x)\,\mathrm{d}x + \lim_{n\to\infty} \int_{\frac{1}{n}}^{\infty} h_n(x)\,\mathrm{d}x$$

$$\begin{aligned}
\lim_{n\to\infty} \int_{-\infty}^{\infty} h_n(x)\,\mathrm{d}x &= \lim_{n\to\infty} 0 + \lim_{n\to\infty} \int_{-\frac{1}{n}}^{\frac{1}{n}} \frac{n}{2}\,\mathrm{d}x + \lim_{n\to\infty} 0 \\
&= \lim_{n\to\infty} \frac{n}{2} \cdot \left[\frac{1}{n} - \left(-\frac{1}{n}\right)\right] \\
&= \lim_{n\to\infty} \frac{n}{2} \cdot \frac{2}{n} \\
&= \lim_{n\to\infty} 1 \\
&= 1
\end{aligned}$$

Então podemos concluir que $\lim_{n\to\infty} h_n(x) = \delta(x,a)$.

Podemos simplificar a notação $\delta(x, a)$ para $\delta(x - a)$, considerando $\delta(x) = \delta(x, 0)$.

Como a singularidade de $\delta(x)$ ocorre em $x = 0$, segue que $\delta(-x) = \delta(x)$. Além disso, a primitiva do delta de Dirac é dada por

$$H(x) = \int_{-\infty}^{x} \delta(x')\,\mathrm{d}x'$$

Note que $H(x) = 0$ se $x < 0$ e $H(x) = 1$ se $x > 0$. Em outras palavras, a função de Heaviside $H(x)$

$$H(x) = \begin{cases} 0 & \text{se } x < 0 \\ \frac{1}{2} & \text{se } x = 0 \\ 1 & \text{se } x > 0 \end{cases}$$

pode ser considerada como uma primitiva de $\delta(x)$.

A primitiva da função de Heaviside é, por sua vez, dada por

$$\begin{aligned} R(x) &= \int_{-\infty}^{x} H(x')\,\mathrm{d}x' \\ &= \int_{-\infty}^{0} H(x')\,\mathrm{d}x' + \int_{0}^{x} H(x')\,\mathrm{d}x' \\ &= 0 + \int_{0}^{x} 1\,\mathrm{d}x' \\ &= x \end{aligned}$$

se $x \geq 0$. E no caso $x < 0$, $R(x) = 0$. Portanto

$$R(x) = \begin{cases} 0 & \text{se } x < 0 \\ x & \text{se } x \geq 0 \end{cases}$$

ou então $R(x) = xH(x)$.

Perceba que $R(x)$ é uma função contínua em todo o domínio, apesar da sua segunda derivada (o delta de Dirac) não ser uma função e apresentar peculiaridades em $x = 0$.

4.2 Funções de Green

Seja L_x um operador diferencial linear aplicável à variável x. A equação diferencial

$$L_x(G(x, s)) = \delta(x - s) \tag{230}$$

é útil para resolver EDOs lineares não-homogêneas do tipo

$$L_x(f(x)) = g(x) \tag{231}$$

Multiplicando (230) por $g(s)$ e integrando em relação à variável s, com $a < x < b$, obtemos

$$\begin{aligned}
\int_a^b L_x(G(x,s))g(s)\,\mathrm{d}s &= \int_a^b \delta(x-s)g(s)\,\mathrm{d}s \\
&= \int_a^b \delta(s-x)g(s)\,\mathrm{d}s \\
&= g(x)
\end{aligned} \tag{232}$$

Por outro lado, o operador L_x pode sair da integral, visto que esta opera em s e o intervalo de integração não depende de x. Então

$$\int_a^b L_x(G(x,s))g(s)\,\mathrm{d}s = L_x\left(\int_a^b G(x,s)g(s)\,\mathrm{d}s\right)$$

ou seja,

$$L_x(f(x)) = g(x) = \int_a^b L_x(G(x,s))g(s)\,\mathrm{d}s = L_x\left(\int_a^b G(x,s)g(s)\,\mathrm{d}s\right) \tag{233}$$

Comparando os argumentos do operador L_x, chegamos à expressão

$$f(x) = \int_a^b G(x,s)g(s)\,\mathrm{d}s + u(x) \tag{234}$$

onde $u(x)$ é qualquer solução da equação diferencial homogênea $L_x u = 0$. A função $g(x)$ a priori já é conhecida e $u(x)$ sabemos calcular. Resta apenas conhecermos a função $G(x,s)$ para obtermos a solução da EDO (231). Ela é denominada de função de Green.

Da equação (230) também é possível afirmar que

$$L_x(G(x,s)) = 0 \quad \text{se} \quad x \neq s \tag{235}$$

4.3 Raízes de soluções de EDOs lineares de segunda ordem

Os próximos teoremas nos auxiliarão a descobrir propriedades de soluções de equações diferenciais lineares de segunda ordem sem precisar resolvê-las.

Teorema 4.1. *Seja $f(x)$ uma função contínua e não nula que satisfaz a equação diferencial ordinária*

$$f''(x) + a_1(x)f'(x) + a_0(x)f(x) = 0 \tag{236}$$

em um intervalo I. Então para cada raiz x_0 de $f(x)$ em I sempre existirá um intervalo I_{x_0} englobando-a tal que ela seja a única raiz em I_{x_0}.

Demonstração. Suponha que $f(x_0) = 0$, com $x_0 \in I$. A solução f será única se impusermos uma outra condição inicial envolvendo a derivada de f em x_0. Se $f'(x_0) = 0$, então a solução será a trivial, $f(x) = 0\ \forall x \in I$. Por hipótese, $f(x)$ não é a função identicamente nula em I, então $f'(x_0) \neq 0$. Isso significa que f está crescendo ou decrescendo em x_0, no que implica na existência de um intervalo I_{x_0} contendo x_0 com tamanho pequeno o suficiente para que x_0 seja a única raiz de f em I_{x_0}. ∎

As raízes x_0 que possuem um intervalo I_{x_0} com as características do teorema (4.1) são denominadas de raízes isoladas.

Teorema 4.2 (Teorema de Sturm da separação)**.** *Se f_1 e f_2 são duas soluções contínuas linearmente independentes da equação diferencial* (236) *em um intervalo I, então as raízes de f_1 são distintas de f_2 e, além disso, todo intervalo (x_0, x_1) entre duas raízes consecutivas x_0 e x_1 de f_2 em I contém exatamente uma raiz de f_1 e vice-versa.*

Demonstração. Dado que f_1 e f_2 são duas soluções linearmente independentes, então o teorema (1.43) nos garante que a wronskiana de f_1 e f_2 nunca se anula em I, ou seja,

$$W(f_1, f_2)(x) = f_1(x)f_2'(x) - f_1'(x)f_2(x) \neq 0 \quad \forall x \in I$$

Note que se f_1 e f_2 tiverem uma mesma raiz, então a wronskiana se anula nesse ponto, no que implica (pelo teorema 1.42) que f_1 e f_2 são linearmente dependentes, o que contradiz a hipótese. Como a wronskiana nunca se anula em I, segue que $W(f_1, f_2)(x)$ possui o mesmo sinal em todo o intervalo I.

Agora suponha que x_0 e x_1 são duas raízes consecutivas de f_2. Então

$$\begin{aligned} W(f_1, f_2)(x_0) &= f_1(x_0)f_2'(x_0) - f_1'(x_0)f_2(x_0) \\ &= f_1(x_0)f_2'(x_0) \\ &\neq 0 \end{aligned}$$

Analogamente, temos

$$f_1(x_1)f_2'(x_1) \neq 0$$

Isso implica que os valores $f_1(x_0)$, $f_2'(x_0)$, $f_1(x_1)$ e $f_2'(x_1)$ são todos não nulos. Como f_2 é uma função contínua em I, o teorema (4.1) garante que x_0 e x_1 são duas raízes isoladas. Isso implica que os sinais de $f_2'(x_0)$ e $f_2'(x_1)$ são opostos entre si, caso contrário haverão pontos em (x_0, x_1) em que $f_2 > 0$ e também $f_2 < 0$, no que implica (pelo teorema do valor intermediário) na existência de uma terceira raiz no intervalo (x_0, x_1), o que contradiz a hipótese. Como a wronskiana possui o mesmo sinal em todo o intervalo I, segue que $f_1(x_0)f_2'(x_0)$ e $f_1(x_1)f_2'(x_1)$ possuem o mesmo sinal, no que implica que $f_1(x_0)$ e $f_1(x_1)$ possuem sinais opostos. Como f_1 é uma função contínua em I, segue que f_1 tem pelo menos uma raiz em (x_0, x_1).

Agora suponha que f_1 tenha duas ou mais raízes em (x_0, x_1). Podemos repetir a argumentação para concluir que f_2 terá pelo menos uma terceira raiz

em (x_0, x_1), o que contradiz a hipótese. Isso significa que f_1 tem apenas uma única raiz em (x_0, x_1). A demonstração de que f_2 tem uma única raiz entre duas consecutivas de f_1 é inteiramente análoga. ■

A contrapositiva do teorema (4.2) resulta no seguinte corolário:

Corolário. *Se f_1 e f_2 são duas soluções contínuas da EDO* (236) *e possuem uma raiz em comum em I, então f_1 e f_2 são linearmente dependentes.*

Teorema 4.3 (Teorema da comparação de Sturm)**.** *Sejam as equações diferenciais*

$$(pf')' + q_1 f = pf'' + p'f' + q_1 f = 0 \tag{237}$$

$$(pg')' + q_2 g = pg'' + p'g' + q_2 g = 0 \tag{238}$$

com f e g sendo soluções não-triviais, contínuas e com derivadas contínuas em um intervalo I e p uma função contínua, diferenciável e não negativa nesse intervalo e que só pode se anular nas extremidades de I. Então se $q_1(x) > q_2(x)$ $\forall x \in I$, então f tem pelo menos uma raiz entre duas consecutivas de g.

Demonstração. Sejam x_0 e x_1 duas raízes consecutivas de g em I. Suponha que f não tenha raízes no intervalo (x_0, x_1). Podemos também supor, sem perda de generalidade, que ambas as funções f e g sejam positivas em (x_0, x_1). Então $g'(x_0) \geq 0$, $g'(x_1) \leq 0$ e

$$p(x_0)W(f,g)(x_0) = p(x_0)[f(x_0)g'(x_0) - f'(x_0)g(x_0)] = p(x_0)f(x_0)g'(x_0) \geq 0$$

Analogamente,

$$p(x_1)W(f,g)(x_1) = p(x_1)f(x_1)g'(x_1) \leq 0$$

Sabemos que $W'(f,g) = fg'' - f''g$ e portanto

$$\begin{aligned} pW'(f,g) &= fpg'' - pf''g \\ &= gp'f' + gq_1 f - fp'g' - fq_2 g \\ &= [q_1 - q_2]fg + p'[gf' - fg'] \\ &= [q_1 - q_2]fg - p'W(f,g) \end{aligned}$$

onde usamos (237) e (238) para substituir pf'' e pg''. Isso implica

$$\begin{aligned} [q_1 - q_2]fg &= pW'(f,g) + p'W(f,g) \\ &= (pW(f,g))' \end{aligned}$$

Mas como $q_1 > q_2$, $f > 0$ e $g \geq 0$ em todo o intervalo I, segue que

$$(pW(f,g))' \geq 0$$

ou seja, a função $pW(f,g)$ é sempre crescente em I, o que contradiz

$$p(x_0)f(x_0)g'(x_0) \geq 0 \quad \text{e} \quad p(x_1)f(x_1)g'(x_1) \leq 0$$

uma vez que p e $W(f,g)$ são funções contínuas e portanto $pW(f,g)$ deveria ser decrescente em I. Há uma ressalva se $(pW(f,g))' = 0$ para todo $x \in I$, mas isso implica $q_1 = q_2\ \forall x \in I$ (pois f e g não são a solução trivial), o que está fora de questão. Esse absurdo nos leva a concluir que f precisa ter pelo menos uma raiz em (x_0, x_1).

A demonstração com ambas as funções f e g não positivas em I é completamente idêntica. E se f e g tiverem sinais opostos, chegaremos à conclusão que

$$p(x_0)f(x_0)g'(x_0) \leq 0 \quad \text{e} \quad p(x_1)f(x_1)g'(x_1) \geq 0$$

enquanto que $(pW(f,g))' \leq 0$, no que acarreta no mesmo absurdo. ■

As aplicações do teorema (4.3) serão usadas nos problemas de Sturm-Liouville, que trataremos adiante.

4.4 O problema de Sturm

Dada uma equação diferencial do tipo

$$f''(x) + a_1(x)f'(x) + a_0(x)f(x) = g(x) \tag{239}$$

podemos definir as funções $p(x) = \exp\left(\int_a^x a_1(x')\,\mathrm{d}x'\right)$, $q(x) = a_0(x)p(x)$ e $h(x) = p(x)g(x)$. E então

$$p'(x) = \left[\frac{d}{dx}\int_a^x a_1(x')\,\mathrm{d}x'\right]\exp\left(\int_a^x a_1(x')\,\mathrm{d}x'\right) = a_1(x)p(x)$$

$$\begin{aligned}
(p(x)f'(x))' + q(x)f(x) &= p'(x)f'(x) + p(x)f''(x) + a_0(x)p(x)f(x) \\
&= p(x)a_1(x)f'(x) + p(x)f''(x) + a_0(x)p(x)f(x) \\
&= p(x)\left[f''(x) + a_1(x)f'(x) + a_0(x)f(x)\right] \\
&= p(x)g(x) \\
&= h(x)
\end{aligned}$$

Isso significa que a EDO (239) é equivalente a equação

$$(p(x)f'(x))' + q(x)f(x) = h(x) \tag{240}$$

O operador que representa essa equação diferencial é o de Sturm-Liouville, conforme vimos no exemplo (1.31):

$$T(f) = (pf')' + qf = h$$

Também vimos que esse operador T é hermitiano no espaço euclidiano $C(a,b)$ das funções reais contínuas no intervalo $[a,b]$ que satisfazem as condições de contorno

$$p(a)f(a) = p(b)f(b) = 0$$

com o produto interno

$$\langle f, g \rangle = \int_a^b f(x)g(x)\,\mathrm{d}x$$

Agora vamos usar um espaço euclidiano mais genérico: $C(a, b)$ é composto pelas funções reais f que são contínuas no intervalo $[a, b]$ e que satisfazem as condições de contorno

$$\alpha_1 f(a) + \beta_1 f'(a) = 0 \quad (\alpha_1; \beta_1) \neq (0; 0) \tag{241}$$

$$\alpha_2 f(b) + \beta_2 f'(b) = 0 \quad (\alpha_2; \beta_2) \neq (0; 0) \tag{242}$$

onde $\alpha_1, \alpha_2, \beta_1$ e β_2 são coeficientes reais. Além disso, o produto interno é dado por

$$\langle f, g \rangle = \int_a^b \overline{f(x)}g(x)\,\mathrm{d}x \tag{243}$$

Nesse novo espaço euclidiano, definimos o operador de Sturm-Liouville L como sendo

$$L = \frac{d}{dx}\left(p(x)\frac{d}{dx}\right) + q(x)$$

onde p, q são funções reais contínuas em $[a, b]$ e p é diferenciável e positiva em $[a, b]$. O operador L também é hermitiano. Para isso, vamos provar que $\langle L(f), g \rangle - \langle f, L(g) \rangle$ é zero para quaisquer funções f e g de $C(a, b)$:

$$\langle L(f), g \rangle - \langle f, L(g) \rangle = \int_a^b (p\overline{f}')'g + \overline{f}qg\,\mathrm{d}x - \int_a^b \overline{f}(pg')' + \overline{f}qg\,\mathrm{d}x$$

$$\begin{aligned}\langle L(f), g \rangle - \langle f, L(g) \rangle &= [\overline{f'}pg - \overline{f}pg']\Big|_a^b + \int_a^b \overline{f}qg - \overline{f}qg - \overline{f'}pg' + \overline{f'}pg'\,\mathrm{d}x \\ &= [\overline{f'}pg - \overline{f}pg']\Big|_a^b\end{aligned} \tag{244}$$

Isto é, precisamos provar

$$p(b)\cdot[\overline{f'(b)}g(b) - \overline{f(b)}g'(b)] - p(a)\cdot[\overline{f'(a)}g(a) - \overline{f(a)}g'(a)] = 0$$

Para isso, é suficiente mostrar que os termos entre parênteses são nulos. Sabendo que os coeficientes das condições de contorno são reais, então é válida a expressão

$$\alpha_1\overline{f(a)} + \beta_1\overline{f'(a)} = 0$$

e então temos o par de equações

$$\alpha_1 g(a) + \beta_1 g'(a) = 0$$
$$\alpha_1\overline{f(a)} + \beta_1\overline{f'(a)} = 0$$

que é equivalente à expressão matricial

$$\begin{bmatrix} g(a) & g'(a) \\ \\ \overline{f(a)} & \overline{f'(a)} \end{bmatrix} \cdot \begin{bmatrix} \alpha_1 \\ \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 0 \\ \\ 0 \end{bmatrix} \tag{245}$$

A matriz coluna dos coeficientes é diferente da matriz nula, pois isso é garantido pela condição de contorno (241). Então, pelo teorema (1.36), a equação matricial homogênea (245) possui soluções além da trivial, o que implica

$$\det \begin{bmatrix} g(a) & g'(a) \\ \\ \overline{f(a)} & \overline{f'(a)} \end{bmatrix} = 0$$

ou seja

$$\overline{f'(a)}g(a) - \overline{f(a)}g'(a) = 0$$

De forma análoga, também podemos encontrar

$$\overline{f'(b)}g(b) - \overline{f(b)}g'(b) = 0$$

o que implica

$$\langle L(f), g\rangle - \langle f, L(g)\rangle = 0$$

Isso significa que, pelo teorema (1.50), podemos encontrar uma base de autovetores ortogonais do operador L em $C(a, b)$ tais que satisfazem

$$Lu = \lambda u$$

uma vez que o espaço de soluções é finito. Além disso, os seus autovalores associados são reais (teorema (1.48)).

Teorema 4.4. *Seja a equação diferencial do tipo*

$$f''(x) + a_1(x)f'(x) + a_0(x)f(x) = g(x)$$

com as condições de contorno

$$\alpha_1 f(a) + \beta_1 f'(a) = c_1 \quad (\alpha_1; \beta_1) \neq (0; 0)$$
$$\alpha_2 f(b) + \beta_2 f'(b) = c_2 \quad (\alpha_2; \beta_2) \neq (0; 0)$$

onde $\alpha_1, \alpha_2, \beta_1, \beta_2, c_1$ e c_2 são coeficientes reais e a_0, a_1 e g são funções reais contínuas em $[a, b]$. A solução existe e é única se e somente se existirem duas funções linearmente independentes u_1 e u_2 tais que são soluções da equação homogênea

$$u_1''(x) + a_1(x)u_1'(x) + a_0(x)u_1(x) = 0$$
$$u_2''(x) + a_1(x)u_2'(x) + a_0(x)u_2(x) = 0$$

e simultaneamente

$$\det \begin{bmatrix} \alpha_1 u_1(a) + \beta_1 u_1'(a) & \alpha_1 u_2(a) + \beta_1 u_2'(a) \\ \alpha_2 u_1(b) + \beta_2 u_1'(b) & \alpha_2 u_2(b) + \beta_2 u_2'(b) \end{bmatrix} \neq 0 \tag{246}$$

Teorema 4.5. *Seja a equação diferencial* (439) *e também as constantes reais* $\alpha_1, \alpha_2, \beta_1$ *e* β_2 *tais que*

$$(\alpha_1; \beta_1) \neq (0; 0) \quad e \quad (\alpha_2; \beta_2) \neq (0; 0)$$

Se v_1 *e* v_2 *são duas soluções linearmente independentes* (439) *definidas no intervalo* $[a, b]$ *e, além disso,* u_1 *e* u_2 *também são duas soluções linearmente independentes da mesma equação no mesmo intervalo* $[a, b]$*, então*

$$\det \begin{bmatrix} \alpha_1 v_1(a) + \beta_1 v_1'(a) & \alpha_1 v_2(a) + \beta_1 v_2'(a) \\ \alpha_2 v_1(b) + \beta_2 v_1'(b) & \alpha_2 v_2(b) + \beta_2 v_2'(b) \end{bmatrix} \neq 0$$

se e somente se

$$\det \begin{bmatrix} \alpha_1 u_1(a) + \beta_1 u_1'(a) & \alpha_1 u_2(a) + \beta_1 u_2'(a) \\ \alpha_2 u_1(b) + \beta_2 u_1'(b) & \alpha_2 u_2(b) + \beta_2 u_2'(b) \end{bmatrix} \neq 0$$

A demonstração do teorema (4.4) está no apêndice (C.1). Além disso, o teorema (4.5) (cuja demonstração também está no apêndice (C.2)) garante que qualquer par de soluções da equação homogênea serve para verificar (246).

Dado esse contexto, o problema de Sturm consiste em encontrar as soluções da equação diferencial não-homogênea

$$Lf = h \tag{247}$$

onde g é uma função contínua em $[a, b]$, a função p do operador L é diferenciável e positiva em $[a, b]$ e f precisa satisfazer as condições de fronteira (241) e (242). Além disso, para quaisquer soluções u_1 e u_2 da equação homogênea $Lu = 0$ é preciso impor

$$\det \begin{bmatrix} \alpha_1 u_1(a) + \beta_1 u_1'(a) & \alpha_1 u_2(a) + \beta_1 u_2'(a) \\ \alpha_2 u_1(b) + \beta_2 u_1'(b) & \alpha_2 u_2(b) + \beta_2 u_2'(b) \end{bmatrix} \neq 0 \tag{248}$$

para que a solução exista e seja única.

4.4.1 A função de Green do operador de Sturm

Vamos primeiro nos concentrar nas soluções da equação homogênea

$$Lu = 0 \tag{249}$$

com as condições de fronteira (241) e (242). Note que se $u(x) \in C(a, b)$, $u(x) \neq 0$ e satisfaz (249) com as condições de contorno, então u é uma autofunção com autovalor 0.

Teorema 4.6. *Se existirem duas soluções* u_1 *e* u_2 *da equação homogênea que obedecem* (246)*, então também existem duas funções* v_1 *e* v_2 *linearmente independentes definidas em* $[a, b]$ *tais que satisfazem* $Lv_1 = Lv_2 = 0$*,*

$$\alpha_1 v_1(a) + \beta_1 v_1'(a) = 0$$
$$\alpha_2 v_2(b) + \beta_2 v_2'(b) = 0$$

e ainda

$$W(v_1, v_2) = \det \begin{bmatrix} v_1(x) & v_2(x) \\ v_1'(x) & v_2'(x) \end{bmatrix} \neq 0 \quad \forall x \in [a, b]$$

$$\alpha_1 v_2(a) + \beta_1 v_2'(a) \neq 0 \tag{250}$$

$$\alpha_2 v_1(b) + \beta_2 v_1'(b) \neq 0 \tag{251}$$

A demonstração do teorema (4.6) está no apêndice (D.1).

Note que as condições (250) e (251) certificam que nem v_1 ou v_2 possam satisfazer ambas as condições de contorno (241) e (242) simultaneamente, de modo que não sejam autofunções de autovalor zero.

Agora, para provarmos que a função $p(x) \cdot W(v_1, v_2)(x)$ é constante e diferente de zero, precisaremos do próximo teorema.

Teorema 4.7 (Identidade de Lagrange)**.** *Dada duas funções u e v diferenciáveis de pelo menos segunda ordem e o operador L de Sturm-Liouville*

$$L = \frac{d}{dx}\left(p\frac{d}{dx}\right) + q$$

então

$$uL(v) - vL(u) = [p(uv' - vu')]' \tag{252}$$

Demonstração.

$$\begin{aligned} [p(uv' - vu')]' &= p'(uv' - vu') + p(uv' - vu')' \\ &= p'uv' - p'vu' + pu'v' + puv'' - pv'u' - pvu'' \\ &= p'uv' - p'vu' + puv'' - pvu'' \end{aligned} \tag{253}$$

$$\begin{aligned} uL(v) - vL(u) &= u(pv')' + uqv - v(pu')' - vqu \\ &= p'uv' + puv'' - p'vu' - pvu'' \end{aligned} \tag{254}$$

Comparando ambas as expressões (253) e (254), concluímos que

$$uL(v) - vL(u) = [p(uv' - vu')]'$$

■

Aplicando o teorema (4.7) às funções v_1 e v_2 do teorema (4.6), obtemos

$$\begin{aligned} [p(v_1 v_2' - v_2 v_1')]' &= v_1 L(v_2) - v_2 L(v_1) \\ [p \cdot \det W(v_1, v_2)]' &= 0 \end{aligned}$$

o que implica

$$p(x) \cdot \det W(v_1, v_2)(x) = c \tag{255}$$

Como a função $p(x)$ é positiva em todo o intervalo $[a, b]$ e o determinante da wronskiana de v_1 e v_2 nunca é zero em $[a, b]$, segue que $c \neq 0$.

Daqui em diante vamos adotar a convenção de que o símbolo $W(v_1, v_2)(x)$ pode ser equivalentemente definido como o determinante da wronskiana:

$$W(v_1, v_2)(x) = v_1(x)v_2'(x) - v_2(x)v_1'(x)$$

assim convenientemente podemos omitir o símbolo de determinante quando mencionarmos a wronskiana. Também vamos considerar $W(x)$ por padrão como sendo a wronskiana das soluções v_1 e v_2 da equação homogênea. Por exemplo, em um caso particular de (255), temos $p(a)W(a) = p(b)W(b)$.

Agora seja $G(x, s)$ a função de Green que satisfaz

$$L_x(G(x,s)) = \frac{\partial}{\partial x}\left(p(x)\frac{\partial}{\partial x}G(x,s)\right) + q(x)G(x,s) = \delta(x-s) \tag{256}$$

com as condições de contorno

$$\alpha_1 G(a,s) + \beta_1 G_x(a,s) = 0 \quad (\alpha_1;\beta_1) \neq (0;0) \tag{257}$$
$$\alpha_2 G(b,s) + \beta_2 G_x(b,s) = 0 \quad (\alpha_2;\beta_2) \neq (0;0) \tag{258}$$

onde L_x é o operador de Sturm-Liouville com as derivadas relacionadas à variável x e G_x é a derivada de G em relação a x. Dessa forma, G é contínua nos intervalos $[a, s)$ e $(s, b]$.

Sabemos que a função $v_1(x)$ satisfaz

$$\alpha_1 v_1(a) + \beta_1 v_1'(a) = 0 \tag{259}$$

Restringindo o domínio de x ao intervalo $[a, s)$, onde $\delta(x - s) = 0$, podemos unir (257) e (259) em uma única expressão matricial:

$$\begin{bmatrix} v_1(a) & v_1'(a) \\ G(a,s) & G_x(a,s) \end{bmatrix} \cdot \begin{bmatrix} \alpha_1 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Como $(\alpha_1; \beta_1) \neq (0; 0)$, segue que essa equação matricial tem mais de uma solução, sugerindo que

$$\det \begin{bmatrix} v_1(a) & v_1'(a) \\ G(a,s) & G_x(a,s) \end{bmatrix} = 0$$

ou equivalentemente

$$W(v_1, G)(a) = 0$$

Como a wronskiana é nula em $x = a$ enquanto que v_1 e G são duas soluções da mesma EDO homogênea no mesmo intervalo, pelo teorema (1.42) a wronskiana deve ser identicamente nula no intervalo $[a, s)$, indicando que, pelo teorema (1.43), $v_1(x)$ e $G(x, s)$ só diferem de uma constante (em relação a x) $d_1(s)$ no intervalo $[a, s)$:

$$G(x,s) = d_1(s)v_1(x) \quad \text{se} \quad x \in [a,s) \tag{260}$$

O raciocínio é análogo para mostrar que

$$G(x,s) = d_2(s)v_2(x) \quad \text{se} \quad x \in (s,b] \tag{261}$$

Ainda não sabemos se G é contínua (ou sequer definida) no ponto $x = s$ (ou na reta $x = s$ se s não for fixado). Ou pior, a priori não sabemos se existe tal função $G(x,s)$ que satisfaça a equação diferencial (256) e simultaneamente as condições de contorno (257) e (258). Porém, se existir e ela for contínua, ela satisfaz os itens do teorema (4.8).

Teorema 4.8 (Função de Green do operador de Sturm-Liouville)**.** *Seja L o operador dado por*

$$L = \frac{d}{dx}\left(p\frac{d}{dx}\right) + q$$

Se existir uma função de Green $G(x,s)$ contínua em $[a,b]\times[a,b]$ associada a L com as condições de contorno

$$\begin{aligned} \alpha_1 G(a,s) + \beta_1 G_x(a,s) = 0 \quad (\alpha_1;\beta_1) \neq (0;0) \\ \alpha_2 G_x(b,s) + \beta_2 G_x(b,s) = 0 \quad (\alpha_2;\beta_2) \neq (0;0) \end{aligned}$$

e que satisfaz

$$L_x(G(x,s)) = \delta(x-s)$$

Então $G(x,s)$ obedecerá as seguintes propriedades:

1. *$G(x,s)$ é simétrica, ou seja, $G(s,x) = G(x,s)$*
2. *$G_x(x,s)$ possui uma descontinuidade em $x = s$ e*

$$\lim_{\epsilon\to 0}[G_x(s+\epsilon,s) - G_x(s-\epsilon,s)] = \frac{1}{p(s)}$$

Demonstração. Suponha que a função de Green enunciada exista e seja contínua. Seja $G(x,t)$ a função de Green que satisfaz

$$L_x G(x,t) = \delta(x-t)$$

e também as mesmas condições de contorno de $G(x,s)$.

Integrando a identidade de Lagrange (252) no intervalo $[a,b]$ e considerando $u = G(x,s)$ e $v = G(x,t)$, obtemos

$$\int_a^b G(x,s)L_x(G(x,t)) - G(x,t)L_x(G(x,s))\,\mathrm{d}x =$$

$$[p(x)(G(x,s)G_x(x,t) - G(x,t)G_x(x,s))]\Big|_a^b \tag{262}$$

No primeiro membro, temos

$$\int_a^b G(x,s)L(G(x,t)) - G(x,t)L(G(x,s))\,\mathrm{d}x = \int_a^b G(x,s)\delta(x-t) - G(x,t)\delta(x-s))\,\mathrm{d}x$$

$$\int_a^b G(x,s)L(G(x,t)) - G(x,t)L(G(x,s))\,\mathrm{d}x = G(t,s) - G(s,t)$$

No segundo membro da equação (262), temos

$$[p(x)(G(x,s)G_x(x,t) - G(x,t)G_x(x,s))]\Big|_a^b =$$
$$p(b)(G(b,s)G_x(b,t) - G(b,t)G_x(b,s)) - p(a)(G(a,s)G_x(a,t) - G(a,t)G_x(a,s))$$

Mas segue das condições de contorno que

$$\begin{bmatrix} G(a,s) & G_x(a,s) \\ G(a,t) & G_x(a,t) \end{bmatrix} \cdot \begin{bmatrix} \alpha_1 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

e

$$\begin{bmatrix} G(b,s) & G_x(b,s) \\ G(b,t) & G_x(b,t) \end{bmatrix} \cdot \begin{bmatrix} \alpha_2 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Mas como $(\alpha_1;\beta_1) \neq (0;0)$ e $(\alpha_2;\beta_2) \neq (0;0)$, segue que o determinante das duas matrizes quadradas é zero, o que implica

$$G(a,s)G_x(a,t) - G(a,t)G_x(a,s) = 0$$
$$G(b,s)G_x(b,t) - G(b,t)G_x(b,s) = 0$$

ou seja,

$$[p(x)(G(x,s)G_x(x,t) - G(x,t)G_x(x,s))]\Big|_a^b = 0$$

Portanto

$$G(t,s) - G(s,t) = 0$$
$$G(t,s) = G(s,t)$$

o que prova o segundo item.

Agora integrando a equação (256) em um intervalo $[s-\epsilon, s+\epsilon]$ em relação a x, com $\epsilon > 0$, obtemos

$$\int_{s-\epsilon}^{s+\epsilon} \frac{\partial}{\partial x}\left(p(x)\frac{\partial}{\partial x}G(x,s)\right) + q(x)G(x,s)\,\mathrm{d}x = \int_{s-\epsilon}^{s+\epsilon} \delta(x-s)\,\mathrm{d}x$$

$$[p(x)G_x(x,s)]\Bigg|_{s-\epsilon}^{s+\epsilon} + \int_{s-\epsilon}^{s+\epsilon} q(x)G(x,s)\,\mathrm{d}x = 1 \tag{263}$$

Agora vamos quebrar a integral em duas e calcular seu valor usando as equações (260) e (261):

$$\begin{aligned}\int_{s-\epsilon}^{s+\epsilon} q(x)G(x,s)\,\mathrm{d}x &= \int_{s-\epsilon}^{s} q(x)G(x,s)\,\mathrm{d}x + \int_{s}^{s+\epsilon} q(x)G(x,s)\,\mathrm{d}x \\ &= d_1(s)\int_{s-\epsilon}^{s} q(x)v_1(x)\,\mathrm{d}x + d_2(s)\int_{s}^{s+\epsilon} q(x)v_2(x)\,\mathrm{d}x\end{aligned}$$

Mas como $q(x), v_1(x)$ e $v_2(x)$ são funções contínuas em $[a,b]$, segue que ambas as integrais tendem a zero quando $\epsilon \to 0^+$ e então

$$\lim_{\epsilon\to 0^+} \int_{s-\epsilon}^{s+\epsilon} q(x)G(x,s)\,\mathrm{d}x = 0$$

Além disso,

$$\begin{aligned}\lim_{\epsilon\to 0^+}[p(x)G_x(x,s)]\Bigg|_{s-\epsilon}^{s+\epsilon} &= \lim_{\epsilon\to 0^+}[p(s+\epsilon)G_x(s+\epsilon,s) - p(s-\epsilon)G_x(s-\epsilon,s)] \\ &= p(s)\lim_{\epsilon\to 0^+}[G_x(s+\epsilon,s) - G_x(s-\epsilon,s)]\end{aligned}$$

uma vez que $p(x)$ é contínua em $[a,b]$. E então o limite da equação (263) se torna

$$p(s)\lim_{\epsilon\to 0^+}[G_x(s+\epsilon,s) - G_x(s-\epsilon,s)] = 1$$

o que significa que o salto da derivada de $G(x,s)$ em $x = s$ é

$$\lim_{\epsilon\to 0^+}[G_x(s+\epsilon,s) - G_x(s-\epsilon,s)] = \frac{1}{p(s)}$$

■

Com o auxílio do teorema (4.8), podemos tentar encontrar uma função de Green $G(x,s)$ que seja contínua e que satisfaça as condições do problema de Sturm.

Dada a condição de continuidade de $G(x,s)$ em $x = s$, então segue das equações (260) e (261) que

$$d_2(s)v_2(s) - d_1(s)v_1(s) = 0$$

E do salto da descontinuidade de $G_x(x,s)$ em $x = s$, obtemos

$$d_2(s)v_2'(s) - d_1(s)v_1'(s) = \frac{1}{p(s)}$$

Em formato matricial, podemos unir ambas as equações:

$$\begin{bmatrix} v_2(s) & -v_1(s) \\ v_2'(s) & -v_1'(s) \end{bmatrix} \cdot \begin{bmatrix} d_2(s) \\ d_1(s) \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{1}{p(s)} \end{bmatrix} \tag{264}$$

Como a matriz quadrada de (264) é a wronskiana $W(v_2, v_1)(s)$ com -1 multiplicando a segunda coluna, segue que seu determinante (usando o axioma da homogeneidade (15) e a propriedade 2 da troca de colunas do teorema (1.30)) é igual ao determinante da wronskiana $W(v_1, v_2)(s) = W(s)$, que é diferente de zero para qualquer que seja $s \in [a, b]$ (conforme já vimos anteriormente). Portanto a solução para $d_1(s)$ e $d_2(s)$ é única. Isso também significa que a matriz é inversível e então podemos obter as soluções multiplicando pela inversa à esquerda:

$$\begin{bmatrix} d_2(s) \\ d_1(s) \end{bmatrix} = \frac{1}{W(s)} \begin{bmatrix} -v_1'(s) & v_1(s) \\ -v_2'(s) & v_2(s) \end{bmatrix} \cdot \begin{bmatrix} 0 \\ \frac{1}{p(s)} \end{bmatrix}$$

que resulta

$$d_2(s) = \frac{v_1(s)}{p(s)W(s)}$$

$$d_1(s) = \frac{v_2(s)}{p(s)W(s)}$$

Uma vez que o termo $p(s)W(s)$ é constante $\forall s \in [a, b]$, podemos substituí-lo por $p(a)W(a)$.

E então encontramos uma função de Green deste problema de Sturm, que é dada por

$$G(x,s) = \begin{cases} \dfrac{v_2(s)v_1(x)}{p(a)W(a)} & \text{se } a \leq x \leq s \leq b \\ \\ \dfrac{v_1(s)v_2(x)}{p(a)W(a)} & \text{se } a \leq s \leq x \leq b \end{cases} \tag{265}$$

Portanto a solução da EDO não-homogênea $L(f) = h$ (247), escrita na forma (234), fica

$$\begin{aligned} f(x) &= \int_a^b G(x,s)h(s)\,\mathrm{d}s \\ &= \int_a^x G(x,s)h(s)\,\mathrm{d}s + \int_x^b G(x,s)h(s)\,\mathrm{d}s \\ &= \frac{v_2(x)}{p(a)W(a)} \int_a^x v_1(s)h(s)\,\mathrm{d}s + \frac{v_1(x)}{p(a)W(a)} \int_x^b v_2(s)h(s)\,\mathrm{d}s \end{aligned}$$

onde $s \leq x$ na primeira integral, $x \leq s$ na segunda e $W(s)$ é a wronskiana entre as funções $v_1(s)$ e $v_2(s)$.

4.4.2 Condições de contorno não-homogêneas

Uma extensão do problema de Sturm é, dado o operador de Sturm L, encontrar a função f que satisfaça à equação diferencial

$$Lf = h$$

e também às condições de contorno

$$\alpha_1 f(a) + \beta_1 f'(a) = \gamma_1 \quad (\alpha_1; \beta_1) \neq (0; 0) \tag{266}$$

$$\alpha_2 f(b) + \beta_2 f'(b) = \gamma_2 \quad (\alpha_2; \beta_2) \neq (0; 0) \tag{267}$$

Para isso, vamos decompor f em outras duas funções u e w

$$f = u + w$$

de forma que

$$Lw = g \qquad Lu = h - g$$

e então

$$Lf = Lu + Lw = h - g + g = h$$

Além disso, $u(x)$ precisa satisfazer as condições de contorno homogêneas (241) e (242) enquanto que $w(x)$ é qualquer função duas vezes diferenciável que satisfaça as condições não-homogêneas (266) e (267). Isso implica que

$$\begin{aligned}
\alpha_1 f(a) + \beta_1 f'(a) &= \alpha_1[u(a) + w(a)] + \beta_1[u'(a) + w'(a)] \\
&= [\alpha_1 u(a) + \beta_1 u'(a)] + [\alpha_1 w(a) + \beta_1 w'(a)] \\
&= \gamma_1
\end{aligned}$$

e

$$\begin{aligned}
\alpha_2 f(b) + \beta_2 f'(b) &= \alpha_2[u(b) + w(b)] + \beta_2[u'(b) + w'(b)] \\
&= [\alpha_2 u(b) + \beta_2 u'(b)] + [\alpha_2 w(b) + \beta_2 w'(b)] \\
&= \gamma_2
\end{aligned}$$

Do problema com as condições de contorno homogêneas, temos

$$u(x) = \int_a^b G(x, s)[h(x) - g(x)] \, \mathrm{d}x$$

ou seja,

$$\begin{aligned}
f(x) &= \int_a^b G(x, s)[h(x) - g(x)] \, \mathrm{d}x + w(x) \\
&= \int_a^b G(x, s) h(x) \, \mathrm{d}x + w(x) - \int_a^b G(x, s) g(x) \, \mathrm{d}x \\
&= \int_a^b G(x, s) h(x) \, \mathrm{d}x + w(x) - \int_a^b G(x, s) L w(x) \, \mathrm{d}x
\end{aligned}$$

4.5 O problema de Sturm-Liouville

Seja L o operador de Sturm-Liouville dado por

$$L = \frac{d}{dx}\left(p(x)\frac{d}{dx}\right) + q(x)$$

onde $p(x) > 0$ e $q(x)$ são funções reais contínuas em $[a, b]$ e $p(x)$ é diferenciável em $[a, b]$. O problema de Sturm-Liouville se trata de encontrar a função $u(x)$ assim como os valores λ que satisfazem

$$L(u(x)) + \lambda r(x)u(x) = 0 \quad \forall x \in [a, b] \tag{268}$$

e as condições de contorno

$$\alpha_1 u(a) + \beta_1 u'(a) = 0 \quad (\alpha_1; \beta_1) \neq (0; 0) \tag{269}$$

$$\alpha_2 u(b) + \beta_2 u'(b) = 0 \quad (\alpha_2; \beta_2) \neq (0; 0) \tag{270}$$

onde $r(x) > 0$ é uma função contínua em $[a, b]$. Ela é denominada de função peso e algumas vezes aparece denotada por $\omega(x)$.

A função $u(x)$ e os valores λ são denominados respectivamente de autofunção e autovalores do problema de Sturm-Liouville, pois se definirmos um outro operador M dado por

$$M = -\frac{1}{r(x)}L$$

segue de (268) que

$$-\frac{1}{r(x)}L(u(x)) - \lambda u(x) = 0$$

$$M(u(x)) = \lambda u(x)$$

e então $u(x)$ é um autovetor de M assim como λ é um autovalor de M.

4.5.1 Propriedades das autofunções e dos autovalores

Seja $C(a, b)$ o espaço vetorial das funções reais contínuas que satisfazem as condições de contorno (269) e (270).

Teorema 4.9. *Seja λ um autovalor do problema de Sturm-Liouville e duas autofunções u_1 e u_2 associadas ao λ, onde $Lu_1 + \lambda ru_1 = Lu_2 + \lambda ru_2 = 0$. Então qualquer combinação linear de u_1 e u_2 também é uma autofunção de autovalor λ. Ou seja, para cada λ as autofunções formam um espaço vetorial. Além disso, u_1 e u_2 são linearmente dependentes. Em outras palavras, esse espaço é de dimensão 1.*

Demonstração. Seja $u_3 = a_1u_1 + a_2u_2$. Então

$$\begin{aligned} Lu_3 + \lambda ru_3 &= L(a_1u_1 + a_2u_2) + \lambda r(a_1u_1 + a_2u_2) \\ &= a_1[L(u_1) + \lambda ru_1] + a_2[L(u_2) + \lambda ru_2] \\ &= 0 \end{aligned}$$

Além disso, pelas condições de contorno:

$$\alpha_1 u_1(a) + \beta_1 u_1'(a) = 0$$
$$\alpha_1 u_2(a) + \beta_1 u_2'(a) = 0$$

No formato matricial, temos

$$\begin{bmatrix} u_1(a) & u_1'(a) \\ u_2(a) & u_2'(a) \end{bmatrix} \cdot \begin{bmatrix} \alpha_1 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Mas como $(\alpha_1; \beta_1) \neq (0; 0)$, segue que o valor do determinante da matriz, $u_1(a)u_2'(a) - u_1'(a)u_2(a)$, é nula. Mas esse também é o determinante da Wronskiana $W(u_1, u_2)(b)$. Como u_1 e u_2 são duas soluções da mesma equação homogênea, segue do teorema (1.42) que $W(u_1, u_2)(x) = 0 \quad \forall x \in [a, b]$. Ou seja, pelo teorema (1.43), u_1 e u_2 são linearmente dependentes, o que implica que a dimensão do espaço vetorial das autofunções associadas ao λ precisa ser 1. ■

A contrapositiva do teorema (4.9) é que se u_1 e u_2 são duas autofunções linearmente independentes, então u_1 e u_2 estão associados a autovalores distintos.

Teorema 4.10. *Os autovalores do problema de Sturm-Liouville são sempre reais.*

Demonstração. Seja $u(x)$ uma autofunção do problema de Sturm-Liouville. Por definição:

$$Lu + \lambda r u = 0$$

$$Lu = -\lambda r u$$

Analogamente, temos

$$\begin{aligned} \overline{Lu} &= -\overline{\lambda r u} \\ &= -r\overline{\lambda u} \end{aligned}$$

uma vez que a função r é real. Sabemos que o operador de Sturm-Liouville L é hermitiano, ou seja

$$\langle u, Lu \rangle - \langle Lu, u \rangle = 0$$

Usando a definição do produto interno (243) para o espaço euclidiano $C(a, b)$, temos

$$\int_a^b \overline{u}Lu - u\overline{Lu}\,\mathrm{d}x = 0$$

Mas também

$$\begin{aligned} \int_a^b \overline{u}Lu - u\overline{Lu}\,\mathrm{d}x &= \int_a^b \overline{u}(-\lambda r u) - u(-r\overline{\lambda u})\,\mathrm{d}x \\ &= (\overline{\lambda} - \lambda)\int_a^b |u|^2 r\,\mathrm{d}x \end{aligned}$$

o que implica

$$(\overline{\lambda} - \lambda) \int_a^b |u|^2 r \, \mathrm{d}x = 0 \tag{271}$$

Como u é uma autofunção, então u não é a função identicamente nula, o que significa que $|u|^2 > 0$. Além disso, a função r é positiva no intervalo $[a, b]$. Então a integral da equação (271) é diferente de zero, o que implica

$$\overline{\lambda} - \lambda = 0$$

O autovalor λ é, portanto, real. ∎

Teorema 4.11. *Qualquer autofunção u do problema de Sturm-Liouville possui semelhanças entre a sua parte real e a parte imaginária, diferindo apenas de uma constante.*

Demonstração. Seja u uma autofunção do problema de Sturm-Liouville. Vamos decompor em outras duas funções:

$$u = u_r + u_i i$$

Substituindo na equação diferencial (268), obtemos

$$Lu_r + \lambda r u_r + [Lu_i + \lambda r u_i]\, i = 0 \quad \forall x \in [a, b]$$

ou seja, ambas as partes real e imaginária dessa equação são nulas:

$$\begin{aligned} Lu_r + \lambda r u_r &= 0 \\ Lu_i + \lambda r u_i &= 0 \end{aligned}$$

Além disso, das condições de contorno temos

$$\begin{aligned} \alpha_1 u(a) + \beta_1 u'(a) &= \alpha_1 u_r(a) + \beta_1 u_r'(a) + [\alpha_1 u_i(a) + \beta_1 u_i'(a)]i \\ 0 &= \alpha_1 u_r(a) + \beta_1 u_r'(a) + [\alpha_1 u_i(a) + \beta_1 u_i'(a)]i \end{aligned}$$

$$\begin{aligned} \alpha_1 u_r(a) + \beta_1 u_r'(a) &= 0 \\ \alpha_1 u_i(a) + \beta_1 u_i'(a) &= 0 \end{aligned}$$

Analogamente,

$$\begin{aligned} \alpha_2 u_r(b) + \beta_2 u_r'(b) &= 0 \\ \alpha_2 u_i(b) + \beta_2 u_i'(b) &= 0 \end{aligned}$$

Isso também significa que u_r e u_i são soluções da mesma equação diferencial com o mesmo autovalor e satisfazem as condições de contorno, o que implica, pelo teorema (4.9), que u_r e u_i são linearmente dependentes. ∎

O teorema (4.11) também significa que não há perda de generalidade se suprimirmos a parte imaginária das autofunções do problema de Sturm-Liouville.

Teorema 4.12 (Ortogonalidade das autofunções)**.** *Se u_1 e u_2 são duas autofunções associadas respectivamente aos autovalores λ_1 e $\lambda_2 \neq \lambda_1$, então*

$$\int_a^b \overline{u_1(x)} u_2(x) r(x)\, \mathrm{d}x = \int_a^b \overline{u_2(x)} u_1(x) r(x)\, \mathrm{d}x = 0$$

Demonstração. Seguindo os mesmos passos da demonstração do teorema (4.10), temos

$$Lu_2 = -\lambda_2 r u_2$$
$$\overline{Lu_1} = -r\overline{\lambda_1 u_1}$$

Da propriedade hermitiana do operador L, temos

$$\langle u_1, Lu_2 \rangle - \langle Lu_1, u_2 \rangle = 0$$

$$\int_a^b \overline{u_1} Lu_2 - u_2 \overline{Lu_1}\, \mathrm{d}x = 0$$

Mas também

$$\begin{aligned}
\int_a^b \overline{u_1} Lu_2 - u_2 \overline{Lu_1}\, \mathrm{d}x &= \int_a^b \overline{u_1}(-\lambda_2 r u_2) - u_2(-r\overline{\lambda_1 u_1})\, \mathrm{d}x \\
&= \int_a^b \lambda_1 r \overline{u_1} u_2 - \lambda_2 r \overline{u_1} u_2\, \mathrm{d}x \\
&= (\lambda_1 - \lambda_2) \int_a^b \overline{u_1} u_2 r\, \mathrm{d}x
\end{aligned}$$

o que implica

$$(\lambda_1 - \lambda_2) \int_a^b \overline{u_1} u_2 r\, \mathrm{d}x = 0$$

Como $\lambda_2 \neq \lambda_1$, segue que

$$\int_a^b \overline{u_1} u_2 r\, \mathrm{d}x = 0$$

Similarmente, temos $\langle u_2, Lu_1 \rangle - \langle Lu_2, u_1 \rangle = 0$, que acarreta

$$\int_a^b \overline{u_2} u_1 r\, \mathrm{d}x = 0$$

■

Note que um outro produto interno pode ser definido de forma que u_1 e u_2 sejam ortogonais no sentido usual da álgebra linear:

$$\langle u_1(x), u_2(x) \rangle = \int_a^b \overline{u_1(x)} u_2(x) r(x)\, \mathrm{d}x \tag{272}$$

usando $r(x) > 0$ como peso. Mas por enquanto continuaremos usando a definição anterior (243).

Teorema 4.13. *Se um problema de Sturm-Liouville satisfaz as condições*

1. $q(x) < 0, \quad \forall x \in [a, b]$
2. $\alpha_1 \beta_1 \leq 0$
3. $\alpha_2 \beta_2 \geq 0$

então todos os autovalores λ *serão positivos.*

Demonstração. Seja u uma autofunção problema de Sturm-Liouville. Do produto interno de u com Lu, temos

$$\begin{aligned} \langle u, Lu \rangle &= \int_a^b \overline{u} Lu \, \mathrm{d}x \\ &= \int_a^b \overline{u}(-\lambda r u) \, \mathrm{d}x \\ &= -\lambda \int_a^b |u|^2 r \, \mathrm{d}x \end{aligned}$$

onde λ é o autovalor associado ao autovetor u.

Mas também, da definição do operador L, temos

$$\begin{aligned} \int_a^b \overline{u} Lu \, \mathrm{d}x &= \int_a^b \overline{u}(pu')' \, \mathrm{d}x + \int_a^b q|u|^2 \, \mathrm{d}x \\ &= [p\overline{u}u'] \Big|_a^b - \int_a^b p|u'|^2 \, \mathrm{d}x + \int_a^b q|u|^2 \, \mathrm{d}x \end{aligned}$$

Igualando as equações e depois multiplicando por -1, chegamos na expressão

$$\lambda \int_a^b |u|^2 r \, \mathrm{d}x = [p\overline{u}u'] \Big|_b^a + \int_a^b p|u'|^2 \, \mathrm{d}x - \int_a^b q|u|^2 \, \mathrm{d}x \tag{273}$$

Das condições de contorno temos

$$\begin{aligned} \alpha_1 u(a) &= -\beta_1 u'(a) \qquad (274) \\ \alpha_1 u(a)\overline{u(a)} &= -\beta_1 u'(a) u(a)\overline{u(a)} \\ \overline{u(a)} u'(a) &= -\frac{\alpha_1}{\beta_1} |u(a)|^2 \end{aligned}$$

De maneira análoga, podemos obter

$$\overline{u(b)} u'(b) = -\frac{\alpha_2}{\beta_2} |u(b)|^2$$

Do primeiro termo da equação (273), temos

$$\begin{aligned}[p\overline{u}u']\Big|_b^a &= p(a)\overline{u(a)}u'(a) - p(b)\overline{u(b)}u'(b) \\ &= \frac{\alpha_2}{\beta_2}|u(b)|^2 p(b) - \frac{\alpha_1}{\beta_1}|u(a)|^2 p(a)\end{aligned}$$

Note que se $\alpha_2\beta_2 \geq 0$ (isto é, possuem o mesmo sinal) e $\alpha_1\beta_1 \leq 0$ (sinais opostos), então

$$[p\overline{u}u']\Big|_b^a \geq 0 \tag{275}$$

visto que $p(x) > 0$ em $[a, b]$. E se $\beta_1 = 0$, da equação (274) temos $\alpha_1 u(a) = 0$, com $\alpha_1 \neq 0$, o que implica $u(a) = 0$. Analogamente, se $\beta_2 = 0$, isso resulta em $u(b) = 0$. De qualquer forma, a inequação (275) é satisfeita.

Sabendo também que o termo $\displaystyle\int_a^b p|u'|^2 \,\mathrm{d}x$ é sempre positivo, portanto da equação (273) ocorre que

$$\lambda \int_a^b |u|^2 r \,\mathrm{d}x \geq \int_a^b -q|u|^2 \,\mathrm{d}x \tag{276}$$

Como $q(x) < 0 \quad \forall x \in [a, b]$, então ambas as integrais da inequação (276) são positivas, no que implica $\lambda > 0$. ∎

4.5.2 A equação integral de Fredholm

A equação do problema de Sturm-Liouville também pode ser escrito na forma

$$Lu = -\lambda r u$$

que constitui um problema de Sturm com $h(x) = -\lambda r(x)u(x)$. Portanto a solução desse problema na forma com a função de Green (234) é

$$u(x) = -\lambda \int_a^b G(x, s)r(s)u(s) \,\mathrm{d}s \tag{277}$$

Se fizermos $K(x, s) = -G(x, s)r(s)$, a expressão (277) fica

$$u(x) = \lambda \int_a^b K(x, s)u(s) \,\mathrm{d}s \tag{278}$$

Essa expressão é denominada de equação integral linear homogênea de Fredholm do segundo tipo.

Podemos definir um operador linear $T : \mathcal{C}(a, b) \mapsto \mathcal{C}(a, b)$ como

$$Tf = -\int_a^b G(x, s)r(s)f(s) \,\mathrm{d}s \tag{279}$$

onde $\mathcal{C}(a,b)$ é o espaço euclidiano das funções reais contínuas em $[a,b]$ e onde também está definido o produto interno

$$\langle f, g\rangle = \int_a^b f(x)g(x)r(x)\,\mathrm{d}x$$

dessa forma, a equação (278) se torna

$$\begin{aligned} u &= \lambda T u \\ T u &= \frac{1}{\lambda} u \end{aligned} \tag{280}$$

Note ainda que

$$Tu = \int_a^b G(x,s)r(s)u(s)\,\mathrm{d}s = \langle G(x,s), u(s)\rangle \tag{281}$$

Para encontrarmos as derivadas de Tu, precisaremos do próximo teorema.

Teorema 4.14. *Se $F(x,t)$ é uma função contínua e diferenciável em x e*

$$f(x) = \int_a^x F(x,t)\,\mathrm{d}t$$

então

$$f'(x) = \int_a^x \frac{\partial F}{\partial x}(x,t)\,\mathrm{d}t + F(x,x)$$

Demonstração. Seja H a função definida pela expressão

$$H(x,y) = \int_a^y F(x,t)\,\mathrm{d}t$$

As derivadas de H são triviais: em relação à variável x usamos a fórmula de Leibniz; já para a variável y usamos o teorema fundamental do cálculo:

$$\frac{\partial H}{\partial x} = \int_a^y \frac{\partial F}{\partial x}(x,t)\,\mathrm{d}t \quad \text{e} \quad \frac{\partial H}{\partial y} = F(x,y)$$

Agora seja $g(x)$ uma função contínua e diferenciável e também $h(x) = H(x, g(x))$. Então, pela regra da cadeia:

$$\begin{aligned} h'(x) &= \frac{\partial H}{\partial x}\cdot\frac{dx}{dx} + \frac{\partial H}{\partial y}\cdot\frac{dy}{dx} \\ &= \frac{\partial H}{\partial x} + \frac{\partial H}{\partial g}\cdot\frac{dg}{dx} \\ &= \int_a^{g(x)} \frac{\partial F}{\partial x}(x,t)\,\mathrm{d}t + F(x,g(x))\cdot g'(x) \end{aligned}$$

Fazendo $g(x) = x$, temos $h(x) = H(x,x) = \int_a^x F(x,t)\,\mathrm{d}t = f(x)$ e portanto

$$f'(x) = \int_a^x \frac{\partial F}{\partial x}(x,t)\,\mathrm{d}t + F(x,x)$$

■

Da equação (279), temos

$$\begin{aligned} -Tf &= \int_a^b G(x,s)r(s)f(s)\,\mathrm{d}s \\ &= \int_a^x G(x,s)r(s)f(s)\,\mathrm{d}s + \int_x^b G(x,s)r(s)f(s)\,\mathrm{d}s \end{aligned}$$

Do teorema (4.14):

$$\begin{aligned} -(Tf)' = \int_a^x G_x(x,s)r(s)f(s)\,\mathrm{d}s + \lim_{\epsilon\to 0^+} G(x,x-\epsilon)r(x-\epsilon)f(x-\epsilon) + \\ \int_x^b G_x(x,s)r(s)f(s)\,\mathrm{d}s - \lim_{\epsilon\to 0^+} G(x,x+\epsilon)r(x+\epsilon)f(x+\epsilon) \end{aligned}$$

onde o limite foi usado devido a descontinuidade da derivada de G em $x = s$. Mas G é contínua nesse ponto, assim como r e f, no que acarreta

$$\begin{aligned} -(Tf)' = \int_a^x G_x(x,s)r(s)f(s)\,\mathrm{d}s + G(x,x)r(x)f(x) + \\ \int_x^b G_x(x,s)r(s)f(s)\,\mathrm{d}s - G(x,x)r(x)f(x) \end{aligned}$$

$$-(Tf)' = \int_a^x G_x(x,s)r(s)f(s)\,\mathrm{d}s + \int_x^b G_x(x,s)r(s)f(s)\,\mathrm{d}s$$

Similarmente, temos

$$\begin{aligned} -(Tf)'' = \int_a^x G_{xx}(x,s)r(s)f(s)\,\mathrm{d}s + \lim_{\epsilon\to 0^+} G_x(x,x-\epsilon)r(x-\epsilon)f(x-\epsilon) + \\ \int_x^b G_{xx}(x,s)r(s)f(s)\,\mathrm{d}s - \lim_{\epsilon\to 0^+} G_x(x,x+\epsilon)r(x+\epsilon)f(x+\epsilon) \end{aligned}$$

$$\begin{aligned} -(Tf)'' = \int_a^x G_{xx}(x,s)r(s)f(s)\,\mathrm{d}s + \int_x^b G_{xx}(x,s)r(s)f(s)\,\mathrm{d}s + \\ r(x)f(x)\cdot \lim_{\epsilon\to 0^+} \left[G_x(x,x-\epsilon) - G_x(x,x+\epsilon)\right] \end{aligned}$$

Usando a expressão (265) da função de Green e a continuidade das funções v_1 e v_2, encontramos que

$$\lim_{\epsilon \to 0^+} [G_x(x, x-\epsilon) - G_x(x, x+\epsilon)] = \frac{1}{p(x)}$$

e portanto

$$-(Tf)'' = \int_a^x G_{xx}(x,s)r(s)f(s)\,\mathrm{d}s + \int_x^b G_{xx}(x,s)r(s)f(s)\,\mathrm{d}s + \frac{r(x)f(x)}{p(x)}$$

Com isso, podemos aplicar o operador L de Sturm-Liouville à função Tf:

$$Lu = (pu')' + qu = pu'' + p'u' + qu$$

$$\begin{aligned}
-L(Tf) = p(x)\cdot\int_a^x G_{xx}(x,s)r(s)f(s)\,\mathrm{d}s + p(x)\cdot\int_x^b G_{xx}(x,s)r(s)f(s)\,\mathrm{d}s + \\
r(x)f(x) + p'(x)\cdot\int_a^x G_x(x,s)r(s)f(s)\,\mathrm{d}s + p'(x)\cdot\int_x^b G_x(x,s)r(s)f(s)\,\mathrm{d}s + \\
q(x)\cdot\int_a^x G(x,s)r(s)f(s)\,\mathrm{d}s + q(x)\cdot\int_x^b G(x,s)r(s)f(s)\,\mathrm{d}s
\end{aligned}$$

$$\begin{aligned}
-L(Tf) = \int_a^x [p(x)G_{xx}(x,s) + p'(x)G_x(x,s) + q(x)G(x,s)]\, r(s)f(s)\,\mathrm{d}s + \\
\int_x^b [p(x)G_{xx}(x,s) + p'(x)G_x(x,s) + q(x)G(x,s)]\, r(s)f(s)\,\mathrm{d}s + \\
r(x)f(x)
\end{aligned}$$

$$\begin{aligned}
-L(Tf) &= \int_a^x L_x(G(x,s))\cdot r(s)f(s)\,\mathrm{d}s + \int_x^b L_x(G(x,s))\cdot r(s)f(s)\,\mathrm{d}s + r(x)f(x) \\
&= r(x)f(x)
\end{aligned}$$

uma vez que a função de Green obedece a relação (235). Ou seja, para qualquer f contínua em $[a, b]$, temos

$$L(Tf) = -r(x)f(x) \tag{282}$$

Além disso, dada duas funções reais f e g do espaço $\mathcal{C}(a, b)$, temos que o produto interno real de f com Tg é

$$\langle f, Tg\rangle = \int_a^b f(x)\left[\int_a^b -G(x,s)r(s)g(s)\,\mathrm{d}s\right] r(x)\,\mathrm{d}x$$

Uma vez que $G(x,s)$ e r são funções contínuas em $[a,b]$, então podemos escrever a última equação como

$$\begin{aligned}\langle f, Tg\rangle &= -\int_a^b \int_a^b G(x,s)f(x)r(x)r(s)g(s)\,\mathrm{d}s\,\mathrm{d}x \\ &= -\int_a^b \int_a^b G(s,x)f(s)r(s)r(x)g(x)\,\mathrm{d}x\,\mathrm{d}s \\ &= \int_a^b g(x)\left[\int_a^b -G(x,s)r(s)f(s)\,\mathrm{d}s\right]r(x)\,\mathrm{d}x \\ &= \langle Tf, g\rangle\end{aligned}$$

uma vez que a superfície de integração é simétrica em x e em s e que $G(s,x) = G(x,s)$. Ou seja, T é hermitiano em $\mathcal{C}(a,b)$.

Para avançarmos, precisaremos apresentar algumas definições que serão úteis.

Definição 4.1. *Dado dois espaços euclidianos V e W, uma transformação linear $T : V \mapsto W$ é denominado limitada se existe algum $M > 0$ global tal que satisfaça*

$$\|Tf\|_W \leq M\|f\|_V \quad \forall f \in V$$

onde $\|Tf\|_W$ é a norma de Tf no espaço W e $\|f\|_V$ é a norma de f no espaço V.

Além disso, definimos a norma de uma transformação linear limitada T como sendo

$$\|T\| = \sup_{\|f\|=1} \|Tf\|, \quad f \in V \tag{283}$$

Segue de (283) que para qualquer $u \in \mathcal{C}(a,b)$, $\|u\| \neq 0$, temos

$$\|Tu\| = \left\|T\left(\frac{u}{\|u\|}\right)\right\| \cdot \|u\| \leq \left(\sup_{\|f\|=1} \|Tf\|\right) \cdot \|u\| \leq \|T\| \cdot \|u\| \tag{284}$$

Por outro lado, podemos verificar de forma trivial que essa desigualdade também se mantém para $u = 0$.

Definição 4.2 (Equicontinuidade uniforme de um conjunto infinito de funções)**.** *Seja $\{f_n(x)\}$ um conjunto de funções reais contínuas definidas em um intervalo I. Ele é dito equicontínuo se para um $\epsilon > 0$ seja possível encontrar um $\delta > 0$ tal que satisfaça*

$$|x-y| < \delta \implies |f_n(x) - f_n(y)| < \epsilon \qquad \forall x,y \in I$$

simultaneamente para todas as escolhas de n.

Definição 4.3 (Função uniformemente contínua)**.** *Uma função real f definida em um intervalo I é dita uniformemente contínua se dado $\epsilon > 0$ existir um $\delta > 0$ tal que satisfaça*

$$|x-y| < \delta \implies |f(x) - f(y)| < \epsilon \quad \forall x,y \in I$$

Note que o δ das definições (4.2) e (4.3) são globais (análogo ao N usado na definição de convergência uniforme de sequência de funções, (2.9)).

Além disso, segue que qualquer função f_n de um conjunto equicontínuo em um intervalo I é uniformemente contínua em I.

Uma função contínua em um intervalo I não necessariamente é uniformemente contínua nesse mesmo intervalo. A primeira diz respeito às variações da imagem ao redor de um ponto do domínio, e ser contínua nesse intervalo significa que em todos os pontos desse intervalo não há descontinuidades. Por outro lado, a continuidade uniforme requer que, além da função ser contínua em I, a função tenha variações da sua imagem limitadas globalmente (e por isso uniforme) em todos subintervalos $[x, y] \subseteq I$ de tamanho limitado por δ. Portanto essa é uma definição mais exigente de continuidade.

Exemplo 4.2 (Função contínua mas não uniformemente contínua)**.** *Seja* $f(x) = \dfrac{1}{x}$ *definida em um intervalo* $(0, a]$. *Sabemos que* f *não tem descontinuidades nesse intervalo. Entretanto, dado um* $\epsilon > 0$ *e um* $\delta > 0$, *escolhendo um* $y < \delta$ *tal que* $0 < y < \dfrac{1}{3\epsilon}$ *e também* $x = y + \dfrac{\delta}{2}$, *temos*

$$|x - y| = \left|y + \frac{\delta}{2} - y\right| = \left|\frac{\delta}{2}\right| < \delta$$

porém

$$\begin{aligned}
|f(x) - f(y)| &= \left|\frac{1}{x} - \frac{1}{y}\right| \\
&= \left|\frac{1}{y + \frac{\delta}{2}} - \frac{1}{y}\right| \\
&= \left|\frac{2}{2y + \delta} - \frac{1}{y}\right| \\
&= \left|\frac{2y - 2y - \delta}{(2y + \delta)y}\right| \\
&= \left|\frac{-\delta}{(2y + \delta)y}\right| \\
&= \frac{\delta}{(2y + \delta)y}
\end{aligned}$$

Mas como $y < \delta$, *temos* $2y + \delta < 3\delta$, *no que implica*

$$|f(x) - f(y)| = \frac{\delta}{(2y + \delta)y} > \frac{\delta}{3\delta y} = \frac{1}{3y} > \epsilon$$

uma vez que $y < \dfrac{1}{3\epsilon} \implies \epsilon < \dfrac{1}{3y}$.

Portanto f *não é uniformemente contínua em* $(0, a]$.

Exemplo 4.3 (Continuidade uniforme e intervalos limitados fechados)**.** *Sejam* $f : \mathbb{R} \mapsto \mathbb{R}, f(x) = x^2$ *e* $g : [-1,1] \mapsto [0,1], g(x) = x^2$*. Embora ambas as funções* f *e* g *sejam contínuas em todo o seu domínio, apenas* g *é uniformemente contínua. Dado um* $\epsilon > 0$*,* $y \in \mathbb{R}$ *e* $x = y + \dfrac{\delta}{2}$ *(que implica* $|x - y| < \delta$*), temos*

$$\begin{aligned}|f(x) - f(y)| &= \left|f\left(y + \frac{\delta}{2}\right) - f(y)\right| \\ &= \left|\left(y + \frac{\delta}{2}\right)^2 - y^2\right| \\ &= \left|y^2 + \frac{2}{2}y\delta + \frac{\delta^2}{4} - y^2\right| \\ &= \left|y\delta + \frac{\delta^2}{4}\right|\end{aligned}$$

Qualquer y *suficientemente grande fará* $|f(x) - f(y)|$ *superar* ϵ*, o que não é o caso para a função* g*.*

Um intervalo limitado fechado é um exemplo de espaço compacto. O exemplo (4.3) é um caso particular do próximo teorema:

Teorema 4.15 (Continuidade uniforme e espaços compactos)**.** *Se* f *é uma função real contínua definida em um espaço compacto* X*, então* f *é uniformemente contínua em* X*.*

A demonstração do teorema (4.15) está no capítulo 7, 5ª seção, página 191, enunciado no Teorema 17 da referência [7].

Como a função de Green é uma função contínua definida em um espaço compacto $[a,b] \times [a,b]$, segue do teorema (4.15) que $G(x,s)$ é uniformemente contínua em $[a,b] \times [a,b]$.

Definição 4.4 (Conjunto uniformemente limitado)**.** *Seja* B *um conjunto de funções reais contínuas definidas em um intervalo* I*. Ele é dito uniformemente limitado se existir algum* $M > 0$ *tal que*

$$|f(x)| \leq M$$

simultaneamente $\forall x \in I$*, e* $\forall f \in B$*.*

Teorema 4.16. *O conjunto das funções* $\{Tu\}$ *com* $u \in \mathcal{C}(a,b)$ *e* $\|u\| \leq 1$ *é equicontínuo e uniformemente limitado.*

Demonstração. Da equação (281), temos

$$|Tu| = |\langle G(x,s), u(s)\rangle| \leq \|G(x,s)\| \cdot \|u\|$$

onde a desigualdade de Cauchy-Schwarz foi usado na última inequação.

Como $G(x,s)$ e $r(x)$ são duas funções contínuas, o valor $\|G(x,s)\|^2 = \int_a^b G^2(x,s)r(x)\,\mathrm{d}x$ existe e é finito, o que significa que existe algum $M > 0$ tal que $\forall x \in [a,b]$:

$$\int_a^b G^2(x,s)r(x)\,\mathrm{d}x \leq M^2(b-a) \implies \|G(x,s)\| \leq M\sqrt{b-a}$$

ou seja,

$$|Tu| \leq M\sqrt{b-a} \cdot \|u\| \quad \forall u \in \mathcal{C}(a,b) \tag{285}$$

e no caso $\|u\| \leq 1$:

$$|Tu| \leq M\sqrt{b-a}$$

o que significa que o conjunto $\{Tu\}$ é uniformemente limitado.

Uma vez que a função de Green $G(x,s)$ é uniformemente contínua no retângulo $[a,b] \times [a,b]$, então dado $\epsilon > 0$ existe um $\delta > 0$ tal que

$$\forall x, y \in [a,b], \quad |x-y| < \delta \implies |G(x,s) - G(y,s)| < \epsilon \quad \forall s \in [a,b]$$

Visto que $r, u \in \mathcal{C}(a,b)$, ambas as funções r e u são limitadas em $[a,b]$ e portanto existe um $M > 0$ tal que $|r(s)u(s)| < M, \quad \forall s \in [a,b]$. Além disso, $|x-y| < \delta$ implica

$$\begin{aligned}
|T(u(x)) - T(u(y))| &= \left| -\int_a^b G(x,s)r(s)u(s)\,\mathrm{d}s + \int_a^b G(y,s)r(s)u(s)\,\mathrm{d}s \right| \\
&= \left| \int_a^b (G(x,s) - G(y,s))r(s)u(s)\,\mathrm{d}s \right| \\
&\leq \int_a^b |G(x,s) - G(y,s)| \cdot |r(s)u(s)|\,\mathrm{d}s \\
&< \int_a^b \epsilon \cdot |r(s)u(s)|\,\mathrm{d}s \\
&< \epsilon \int_a^b M\,\mathrm{d}s \\
&< \epsilon M(b-a)
\end{aligned}$$

com $\epsilon > 0$ arbitrário. Logo o conjunto $\{Tu\}$ é equicontínuo.

■

Teorema 4.17. *Dado $u \in \mathcal{C}(a,b)$ com $\|u\| = 1$ e o operador T definido pela expressão* (279)*, a norma de T pode ser dada por*

$$\|T\| = \sup_{\|u\|=1} |\langle Tu, u\rangle|$$

A demonstração desse teorema está no apêndice (E.1).

Teorema 4.18 (Teorema de Ascolí-Arzela)**.** *Se um conjunto de funções reais $f : [a,b] \mapsto \mathbb{R}$ é uniformemente limitado e equicontínuo em $[a,b]$, então existe algum subconjunto $\{f_n(x)\}$ cuja sequência converge uniformemente para uma função contínua $f(x)$.*

O teorema anterior (4.18) é na verdade um corolário do teorema de Ascolí-Arzela, mas devido à sua grande importância e aplicação, em alguns textos também leva o nome de teorema. A demonstração do teorema e do seu corolário (4.18) estão nas páginas 30 a 33 da referência [8].

Teorema 4.19. *Ou $\|T\|$ ou $-\|T\|$ são autovalores de T.*

A demonstração do teorema (4.19) está no apêndice (F.1).

4.5.3 A relação de completude das autofunções

Seja φ_0 a autofunção associada ao autovalor $\mu_0 = \|T\|$ tal que

$$T\varphi_0 = \mu_0\varphi_0 \tag{286}$$

e seja a função $\psi_0 = \dfrac{\varphi_0}{\|\varphi_0\|}$. Defina também

$$G_1(x,s) = G(x,s) + \mu_0\psi_0(x)\psi_0(s)$$

e o operador $T_1 : \mathcal{C}(a,b) \mapsto \mathcal{C}(a,b)$:

$$\begin{aligned}
T_1(u(x)) &= -\int_a^b G_1(x,s)r(s)u(s)\,\mathrm{d}s \\
&= -\int_a^b G(x,s)r(s)u(s) + \mu_0\psi_0(x)\psi_0(s)u(s)r(s)\,\mathrm{d}s \\
&= -\int_a^b G(x,s)r(s)u(s)\,\mathrm{d}s - \mu_0\psi_0(x)\int_a^b \psi_0(s)u(s)r(s)\,\mathrm{d}s \\
&= Tu - \mu_0\psi_0(x)\langle\psi_0, u\rangle
\end{aligned} \tag{287}$$

A função $G_1(x,s)$ também é contínua em $[a,b] \times [a,b]$ e simétrica nesse espaço, assim como $G(x,s)$. Ou seja, T_1 herda a característica hermitiana e os teoremas (4.16), (4.17) e (4.19) também se aplicam para T_1. Se $\|T_1\| \neq 0$, então podemos definir μ_1 como sendo

$$|\mu_1| = \|T_1\| = \sup_{\|u\|=1} |\langle T_1 u, u\rangle|, \quad u \in C(a,b)$$

com $\mu_1 \neq 0$ e que, pelo teorema (4.19), existe uma autofunção φ_1 tal que

$$T_1\varphi_1 = \mu_1\varphi_1$$

Agora seja $\psi_1 = \dfrac{\varphi_1}{\|\varphi_1\|}$. Então $\|\psi_1\|^2 = \langle \psi_0, \psi_0 \rangle = 1$ e

$$\begin{aligned}
\langle T_1 u, \psi_0 \rangle &= \langle Tu - \mu_0 \psi_0(x) \langle \psi_0, u \rangle, \psi_0 \rangle \\
&= \langle Tu, \psi_0 \rangle - \langle \mu_0 \psi_0(x) \langle \psi_0, u \rangle, \psi_0 \rangle \\
&= \langle u, T\psi_0 \rangle - \mu_0 \langle \psi_0, u \rangle \cdot \langle \psi_0, \psi_0 \rangle \\
&= \mu_0 \langle u, \psi_0 \rangle - \mu_0 \langle u, \psi_0 \rangle \\
&= 0
\end{aligned}$$

onde foi usada a relação (286) e o fato do operador T ser hermitiano. Ou seja, a imagem de T_1 é ortogonal à função ψ_0. Daí também temos

$$0 = \langle T_1 \psi_1, \psi_0 \rangle = \mu_1 \langle \psi_1, \psi_0 \rangle \implies \langle \psi_1, \psi_0 \rangle = 0$$

Isso também significa que, da equação (287), temos

$$\begin{aligned}
T_1 \psi_1 &= T\psi_1 - \psi_0(x) \cdot \mu_0 \langle \psi_0, \psi_1 \rangle \\
\mu_1 \psi_1 &= T\psi_1
\end{aligned}$$

ou seja, ψ_1 também é uma autofunção de T. Além disso:

$$\|T\psi_1\| = \|\mu_1 \psi_1\| = |\mu_1| \cdot \|\psi_1\| = |\mu_1|$$

e também $\|T\psi_1\| \leq \|T\| = |\mu_0|$, no que implica

$$|\mu_1| \leq |\mu_0|$$

Analogamente podemos definir uma função $G_2(x,s)$ como

$$\begin{aligned}
G_2(x,s) &= G_1(x,s) + \mu_1 \psi_1(x) \psi_1(s) \\
&= G(x,s) + \sum_{n=0}^{1} \mu_n \psi_n(x) \psi_n(s)
\end{aligned}$$

e um operador T_2

$$\begin{aligned}
T_2(u(x)) &= T_1 u - \mu_1 \psi_1(x) \langle \psi_1, u \rangle \\
&= Tu - \sum_{n=0}^{1} \mu_n \psi_n(x) \langle \psi_n, u \rangle
\end{aligned}$$

para descobrir a existência de um terceiro autovalor μ_2 de T de onde

$$|\mu_2| = \sup_{\|u\|=1} |\langle T_2 u, u \rangle|$$

e

$$|\mu_2| \leq |\mu_1| \leq |\mu_0|$$

assim como sua respectiva autofunção associada ψ_2 de norma 1 que é ortogonal a ambos ψ_1 e ψ_0. Podemos repetir esse processo até encontrarmos algum operador T_m tal que $\|T_m\| = 0$, no que implicaria $|\mu_m| = 0$ e $\mu_m = 0$. Ou seja, ψ_m seria a última autofunção de T na lista decrescente dos módulos dos autovalores. Entretanto, se esse for o caso, temos

$$
\begin{aligned}
LT_m u &= LTu - \sum_{n=0}^{m-1} \mu_n \langle \psi_n, u \rangle L(\psi_n(x)) \\
L(\mu_m u) &= -ru - \sum_{n=0}^{m-1} \langle \psi_n, u \rangle L(\mu_n \psi_n(x)) \\
0 &= -ru - \sum_{n=0}^{m-1} \langle \psi_n, u \rangle LT\psi_n(x) \\
0 &= -ru + \sum_{n=0}^{m-1} \langle \psi_n, u \rangle r\psi_n(x)
\end{aligned}
$$

onde foi usada a relação (282). Como $r(x) > 0$ em $[a, b]$, podemos dividi-lo em ambos os lados da equação e então chegarmos na expressão

$$
u(x) = \sum_{n=0}^{m-1} \langle \psi_n, u \rangle \psi_n(x) \tag{288}
$$

Note que (288) se assemelha à equação (31) da projeção de um elemento de um espaço euclidiano em um subespaço. Isto é, a equação (288) nos diz que qualquer u de $\mathcal{C}(a, b)$ pode ser escrito como combinação linear de um conjunto ortonormal *finito* $\{\psi_0, \ldots, \psi_{m-1}\}$. Entretanto $\mathcal{C}(a, b)$ é de dimensão infinita e isso, portanto, não pode ser verdade, o que implica que não é possível encontrar um operador T_m tal que $\|T_m\| = 0$. Em outras palavras, podemos obter infinitos autovalores cujo módulo é diferente de zero. Isso culmina no teorema a seguir:

Teorema 4.20. *O operador T definido pela expressão* (279) *possui infinitas autofunções ortonormais em $\mathcal{C}(a, b)$.*

Apesar de podermos encontrar autovalores com módulos cada vez menores mas nunca nulos, ainda não há certeza se 0 é o ínfimo desses módulos. Se esse for o caso, o procedimento usado para chegarmos na equação (288) dá indícios de que

$$
u(x) = \sum_{n=0}^{\infty} \langle \psi_n, u \rangle \psi_n(x) \quad \forall u \in C(a, b) \tag{289}
$$

Para podermos provar essa expressão, precisaremos nos ocupar do próximo teorema.

Teorema 4.21. *Dada qualquer função $u \in \mathcal{C}(a, b)$, a série infinita $\sum \langle \psi_n, u \rangle \psi_n$ converge uniformemente em $[a, b]$ para uma função contínua, que é Tu.*

Demonstração. Da equação (281), temos que $\forall n \leq 0$:

$$\langle G(x,s), \psi_n \rangle = T\psi_n(x) = \mu_n \psi_n(x)$$

Com isso, a desigualdade de Bessel (1.14) nos diz que

$$\begin{aligned} \sum_{n=0}^{m} |\langle G(x,s), \psi_n \rangle|^2 &\leq \|G(x,s)\|^2 \\ \sum_{n=0}^{m} \mu_n^2 |\psi_n(x)|^2 &\leq \int_a^b |G(x,s)|^2 r(s)\,\mathrm{d}s \end{aligned}$$

para todo x em $[a,b]$ e para todo $m \geq 0$. Multiplicando ambos os lados por $r(x)$ e integrando em relação a x no intervalo $[a,b]$, obtemos

$$\begin{aligned} \int_a^b \sum_{n=0}^{m} \mu_n^2 |\psi_n(x)|^2 r(x)\,\mathrm{d}x &\leq \int_a^b \int_a^b |G(x,s)|^2 r(s) r(x)\,\mathrm{d}s\,\mathrm{d}x \\ \sum_{n=0}^{m} \mu_n^2 \int_a^b |\psi_n(x)|^2 r(x)\,\mathrm{d}x &\leq \int_a^b \int_a^b M^2 r(s) r(x)\,\mathrm{d}s\,\mathrm{d}x \\ \sum_{n=0}^{m} \mu_n^2 \|\psi_n(x)\|^2 &\leq M^2 R^2 (b-a)^2 \\ \sum_{n=0}^{m} \mu_n^2 &\leq M^2 R^2 (b-a)^2 \end{aligned}$$

onde $M = \max\{G(x,s),\ (x,s) \in [a,b] \times [a,b]\}$ e R é um valor tal que $R > r(x)$, $\forall x \in [a,b]$, uma vez que $r(x) > 0$ é contínua e portanto limitada. Como a expressão é válida $\forall m \geq 0$, podemos fazer $m \to \infty$ e verificar que a série infinita de termos positivos

$$\sum_{n=0}^{\infty} \mu_n^2$$

é limitada superiormente por $M^2 R^2 (b-a)^2$. Ou seja, o teorema (2.7) garante que essa série converge. E por essa causa, o teorema (2.6) garante que, como condição necessária

$$\begin{aligned} \lim_{n \to \infty} \mu_n^2 &= 0 \\ \lim_{n \to \infty} |\mu_n| &= 0 \end{aligned} \tag{290}$$

Agora, da equação (284), temos

$$\begin{aligned} \left\| Tu - \sum_{n=0}^{m-1} \mu_n \langle \psi_n, u \rangle \psi_n(x) \right\| &= \|T_m u\| \\ &\leq \|T_m\| \cdot \|u\| \\ &\leq |\mu_m| \cdot \|u\| \end{aligned}$$

e portanto, devido à relação (290):

$$\left\| Tu - \sum_{n=0}^{m-1} \mu_n \langle \psi_n, u \rangle \psi_n(x) \right\| \to 0 \quad \text{quando} \quad m \to \infty \tag{291}$$

Além disso, da linearidade de T, com $m > l$, temos

$$T\left(\sum_{n=j}^{m} \langle u, \psi_n \rangle \psi_n\right) = \sum_{n=j}^{m} \mu_n \langle u, \psi_n \rangle \psi_n$$

mas da equação (285):

$$\left| T\left(\sum_{n=j}^{m} \langle u, \psi_n \rangle \psi_n\right) \right| \leq M\sqrt{b-a} \cdot \left\| \sum_{n=j}^{m} \langle u, \psi_n \rangle \psi_n \right\|$$
$$\left| \sum_{n=j}^{m} \mu_n \langle u, \psi_n \rangle \psi_n \right| \leq M\sqrt{b-a} \cdot \left\| \sum_{n=j}^{m} \langle u, \psi_n \rangle \psi_n \right\|$$

Já a norma da última expressão pode ser simplificada usando a equação (22):

$$\left\| \sum_{n=j}^{m} \langle u, \psi_n \rangle \psi_n \right\| = \left(\sum_{n=j}^{m} |\langle u, \psi_n \rangle|^2\right)^{1/2}$$

e portanto

$$\left| \sum_{n=j}^{m} \mu_n \langle u, \psi_n \rangle \psi_n \right| \leq M\sqrt{b-a} \cdot \left(\sum_{n=j}^{m} |\langle u, \psi_n \rangle|^2\right)^{1/2}$$

Mas a desigualdade de Bessel (1.14) afirma que

$$\sum_{n=0}^{\infty} |\langle u, \psi_n \rangle|^2 \leq \|u\|^2 \in \mathbb{R} \tag{292}$$

e consequentemente, fazendo $m \to +\infty$, existe um J tal que $\forall j > J$

$$\left(\sum_{n=j}^{\infty} |\langle u, \psi_n \rangle|^2\right)^{1/2} < \epsilon_0$$

Isso significa que

$$j > J \implies \left| \sum_{n=j}^{\infty} \mu_n \langle u, \psi_n \rangle \psi_n \right| < \epsilon \tag{293}$$

onde $\epsilon = M\epsilon_0\sqrt{b-a}$. Agora vamos denotar a soma parcial $s_j(x)$ como sendo

$$s_j(x) = \sum_{n=j}^{\infty} \mu_n \langle u, \psi_n \rangle \psi_n$$

Então podemos reescrever a proposição (293) como sendo

$$j > J \implies |s_j(x) - s(x)| < \epsilon \tag{294}$$

onde $s(x) = 0$. Essa é exatamente a definição de convergência uniforme (2.9). E como da equação (291) temos

$$Tu - \sum_{n=0}^{m-1} \mu_n \langle \psi_n, u \rangle \psi_n(x) = \sum_{n=j}^{\infty} \mu_n \langle u, \psi_n \rangle \psi_n$$

com $j = m$ e o segundo membro converge uniformemente, segue que

$$\sum_{n=0}^{\infty} \mu_n \langle \psi_n, u \rangle \psi_n(x) \tag{295}$$

converge uniformemente para uma função contínua, que é $(Tu)(x)$. ■

Mas isso também implica

$$\begin{aligned} LTu &= \sum_{n=0}^{\infty} \langle \psi_n, u \rangle L(\mu_n \psi_n(x)) \\ -ru &= \sum_{n=0}^{\infty} \langle \psi_n, u \rangle LT\psi_n(x) \\ -ru &= -\sum_{n=0}^{\infty} \langle \psi_n, u \rangle r\psi_n(x) \end{aligned}$$

Dividindo ambos os lados por $-r$, chegamos à expressão

$$u(x) = \sum_{n=0}^{\infty} \langle \psi_n, u \rangle \psi_n(x) \tag{296}$$

Da equação (280), vimos que

$$Tu = \frac{1}{\lambda} u$$

e portanto os autovalores do problema de Sturm-Liouville estão relacionados com os autovalores de T por

$$\mu_n = \frac{1}{\lambda_n}$$

Como $|\mu_n|$ tende a zero quando $n \to \infty$, segue que $|\lambda_n| \to \infty$ quando $n \to \infty$. Ou seja, os autovalores do problema de Sturm-Liouville podem ser rearranjados da seguinte forma:

$$\lambda_0 < \lambda_1 < \lambda_2 < \dots$$

enquanto que as autofunções do problema de Sturm-Liouville são as mesmas autofunções de T.

Por esse motivo, a equação (296) é denominada de expansão em autofunções do problema de Sturm-Liouville.

Além disso, segue do teorema (4.21) que

$$\begin{aligned}
(Tu)(x) &= \sum_{n=0}^{\infty} \frac{1}{\lambda_n} \langle \psi_n, u \rangle \, \psi_n(x) \\
&= \sum_{n=0}^{\infty} \frac{1}{\lambda_n} \cdot \int_a^b \psi_n(s)\psi_n(x) r(s) u(s) \, \mathrm{d}s \\
&= \int_a^b \left[\sum_{n=0}^{\infty} \frac{\psi_n(s)\psi_n(x)}{\lambda_n} \right] r(s) u(s) \, \mathrm{d}s
\end{aligned}$$

Mas também, por definição

$$(Tu)(x) = -\int_a^b G(x,s) r(s) u(s) \, \mathrm{d}s$$

Comparando ambas as expressões e sabendo que a função de Green $G(x,s)$ é única dada as condições de contorno e a inexistência de autovalores nulos, chegamos à conclusão que

$$G(x,s) = -\sum_{n=0}^{\infty} \frac{\psi_n(s)\psi_n(x)}{\lambda_n} \tag{297}$$

Isso significa que podemos obter a função de Green por meio das autofunções de T, que são as mesmas autofunções do problema de Sturm-Liouville. A equação (297) também é conhecida como fórmula de Mercer.

A relação de completude (296) também é válida no espaço $\mathcal{L}^2(a,b)$ das funções quadrado integráveis no intervalo $[a,b]$. A demonstração está no teorema 2.28 da referência [9].

4.5.4 O problema singular de Sturm-Liouville

Em alguns problemas, a função $p(x)$ pode se anular em $x = a$, em $x = b$ ou ambos. Esse é a situação das equações diferenciais de Legendre (298) e de Bessel

esférica (299)

$$\frac{d}{dx}\left((1-x^2)\frac{df}{dx}\right)+\lambda f(x)=0 \quad x\in[-1,1] \tag{298}$$

$$\frac{d}{dx}\left(x^2\frac{df}{dx}\right)+(\lambda x^2-l(l+1))f(x)=0 \quad x\in[0,b] \tag{299}$$

$$\tag{300}$$

Nesse caso, temos que da equação (244):

$$\langle L(f),g\rangle-\langle f,L(g)\rangle=[\overline{f'}pg-\overline{f}pg']\Big|_a^b=0 \tag{301}$$

desde que os limites de f e g existam e sejam finitos em $x=a$ e em $x=b$. No caso particular da função de Bessel esférica, é requisito que $f(b)=0$. Nessas circunstâncias, o operador de Sturm-Liouville continua hermitiano em $C(a,b)$.

Além disso, também há a possibilidade do intervalo trabalhado ser infinito, como por exemplo todo o eixo real. Esse é o caso da equação diferencial do oscilador harmônico quântico (302) e do átomo de hidrogênio (303):

$$f''(x)+[\lambda-x^2]f(x)=0 \quad x\in(-\infty,+\infty) \tag{302}$$

$$\frac{d}{dx}\left(x^2\frac{df}{dx}\right)+[\lambda x-x^2-l(l+1)]f(x)=0 \quad x\in[0,\infty) \tag{303}$$

Nesses casos, uma condição suficiente para que a expressão (301) seja válida é que $\sqrt{r(x)}u(x)\to 0$ quando $|x|\to\infty$.

Um tratamento rigoroso sobre esses e outros resultados da teoria de Sturm-Liouville para os problemas singulares pode ser visto na referência [10].

5 O oscilador harmônico quântico

Vimos na seção (3.3) que se $U(x)$ é uma função analítica com mínimo local em $x = x_0$ e que representa o potencial de um sistema físico, então podemos fazer uma aproximação de Taylor centrada em x_0 tal que

$$U(x) \approx \frac{1}{2}U''(x_0) \cdot (x - x_0)^2 \tag{304}$$

para todo x próximo o suficiente de x_0. Esse é o potencial do oscilador harmônico clássico, onde $U''(x_0)$ representa a constante de elasticidade da mola.

O que difere o oscilador clássico do quântico é que esse potencial é aplicado em leis de movimento completamente diferentes: no clássico, o potencial é usado com a lei de Newton (ou equivalentemente as leis da mecânica lagrangiana ou hamiltoniana); já no problema quântico, é usado a equação de Schrödinger . As equações diferenciais que resultam são diferentes e, portanto, as soluções desses dois problemas também são distintas uma da outra.

Ao definir a frequência angular como

$$\omega = \sqrt{\frac{k}{m}}$$

onde k é a constante de elasticidade da mola e m é a massa constante atrelada a ela, obtemos

$$k = m\omega^2$$

e portanto podemos reescrever o potencial como

$$U(x) = \frac{1}{2}m\omega^2(x - x_0)^2$$

Colocando esse potencial na equação de Schrödinger independente do tempo

$$-\frac{\hbar^2}{2m}\Psi''(x) + [U(x) - E]\Psi(x) = 0 \tag{305}$$

teremos a equação diferencial

$$\Psi''(x - x_0) + \left[\frac{2mE}{\hbar^2} - \frac{m^2}{\hbar^2}\omega^2(x - x_0)^2\right]\Psi(x - x_0) = 0$$

Substituindo x por uma variável adimensional $\xi = \sqrt{\frac{m\omega}{\hbar}}(x - x_0)$ e usando a regra da cadeia, a equação diferencial anterior se torna

$$\Psi''(\xi) + \left[K - \xi^2\right]\Psi(\xi) = 0 \quad \text{onde} \quad K = \frac{2E}{\hbar\omega} \tag{306}$$

Quando ξ^2 é muito grande comparado à constante K, podemos desprezar K e então ficamos com

$$\Psi''(x) = \xi^2\Psi(\xi) \tag{307}$$

Quando ξ^2 é muito maior que K, vimos que a solução da equação se comporta como $e^{\lambda\xi^2}$, pois

$$\frac{d^2}{d\xi^2}e^{\lambda\xi^2} = \left(2\lambda + 4\lambda^2\xi^2\right)e^{\lambda\xi^2}$$

Mas como ξ^2 é muito grande, podemos desprezar o termo 2λ e concluir que

$$\frac{d^2}{d\xi^2}e^{\lambda\xi^2} \approx 4\lambda^2\xi^2 e^{\lambda\xi^2}$$

Para se assemelhar à equação diferencial (307), fazemos $4\lambda^2 = 1$, ou então

$$\lambda = \pm\frac{1}{2} \qquad \lambda_0 = -\frac{1}{2} \qquad \lambda_1 = \frac{1}{2}$$

A solução da equação diferencial (306) que estamos procurando precisa, então, convergir para a solução

$$\Psi(\xi) \approx Ae^{\lambda_0\xi^2} + Be^{\lambda_1\xi^2}$$

quando $\xi^2 >> K$. Para que a solução $\Psi(\xi)$ seja quadrado integrável de $-\infty$ a ∞, é condição necessária que $\Psi(\xi)$ tenda a zero em $+\infty$ ou $-\infty$, o que só acontece se fizermos $B = 0$.

Isso nos diz que

$$\Psi(\xi) = f(\xi)e^{\lambda_0\xi^2} = f(\xi)e^{-\xi^2/2} \tag{308}$$

para alguma função $f(\xi)$. Derivando (308) duas vezes, obtemos

$$\begin{aligned}\Psi''(\xi) &= 2\lambda_0 f(\xi)e^{\lambda_0\xi^2} + 4\lambda_0^2\xi^2 f(\xi)e^{\lambda_0\xi^2} + 2\cdot 2\lambda_0\xi f'(\xi)e^{\lambda_0\xi^2} + f''(\xi)e^{\lambda_0\xi^2} \\ &= \left[(4\lambda_0^2\xi^2 + 2\lambda_0)f(\xi) + 2\cdot 2\lambda_0\xi f'(\xi) + f''(\xi)\right]e^{\lambda_0\xi^2}\end{aligned}$$

Substituindo $\Psi''(\xi)$ e $\Psi(\xi)$ em (306), obtemos

$$\left[(4\lambda_0^2\xi^2 + 2\lambda_0)f(\xi) + 2\cdot 2\lambda_0\xi f'(\xi) + f''(\xi)\right]e^{\lambda_0\xi^2} = [\xi^2 - K]f(x)e^{\lambda_0\xi^2}$$

$$(4\lambda_0^2\xi^2 + 2\lambda_0)f(\xi) + 2\cdot 2\lambda_0\xi f'(\xi) + f''(\xi) = [\xi - K]f(\xi)$$

Logo

$$\begin{aligned}0 &= f''(\xi) + 4\lambda_0\xi f'(\xi) + [4\lambda_0^2\xi^2 - \xi^2 + K + 2\lambda_0]f(\xi) \\ &= f''(\xi) + 4\lambda_0\xi f'(\xi) + \left[4\frac{1}{4}\xi^2 - \xi^2 + K + 2\lambda_0\right]f(\xi) \\ &= f''(\xi) + 4\lambda_0\xi f'(\xi) + (K + 2\lambda_0)f(\xi) \qquad (309)\end{aligned}$$

Sabemos que

$$4\lambda_0 = -4\cdot\frac{1}{2} = -2 \quad \text{e} \quad 2\lambda_0 = -1$$

e portanto a equação diferencial se torna

$$f''(\xi) - 2\xi f'(\xi) + (K - 1)f(\xi) = 0 \tag{310}$$

Dado que as funções coeficientes -2ξ e $K-1$ são analíticas em $\mathbb{R}$, segue que existem duas soluções analíticas linearmente independentes da equação diferencial (310). Ou seja, podemos escrever

$$f(\xi) = \sum_{n=0}^{\infty} a_n \xi^n \tag{311}$$

Derivando a última equação uma, duas vezes e substituindo em (310), obtemos

$$\sum_{n=0}^{\infty} [(n+2)(n+1)a_{n+2} + (-2n+K-1)a_n] \cdot \xi^n = 0$$

Visto que os monômios ξ^n são linearmente independentes e o somatório é zero, então o termo entre colchetes precisa ser nulo para todo $n \geq 0$:

$$(n+2)(n+1)a_{n+2} + (-2n+K-1)a_n = 0$$

no que resulta

$$a_{n+2} = \frac{(2n+1-K)a_n}{(n+2)(n+1)} \tag{312}$$

que é a relação de recorrência para os termos a_n.

Usando o princípio da indução, podemos provar que os coeficientes a_n da solução são dados por

$$a_{2n} = \frac{a_0}{(2n)!} \cdot \prod_{j=0}^{n-1} ((4j+1)-K) \quad \text{se} \quad n \geq 1 \tag{313}$$

e

$$a_{2n+1} = \frac{a_1}{(2n+1)!} \cdot \prod_{j=0}^{n-1} ((4j+3)-K) \quad \text{se} \quad n \geq 1 \tag{314}$$

Como vimos anteriormente, essa sequência pode ser aproximada por uma exponencial e^{ξ^2} de maneira que

$$\Psi(\xi) = f(\xi)e^{-\xi^2/2} \approx Ce^{\xi^2}e^{-\xi^2/2} \approx Ce^{\xi^2/2}$$

que produz uma função que não é normalizável.

Porém $\Psi(\xi)$ precisa ser quadrado integrável, e para isso acontecer é preciso que o somatório em (311) tenha finitos termos, uma vez que os coeficientes a_n fazem a função explodir em $+\infty$ e em $-\infty$. Daí segue que é necessário que $2N+1-K=0$ para algum $N>0$ para que os termos seguintes se anulem, além de impor $a_0=0$ ou $a_1=0$ se N for ímpar ou par respectivamente. Isso pode ser visto na seguinte equação:

$$f(\xi) = \sum_{n=0}^{\infty} a_n \xi^n = \sum_{n=0}^{\infty} a_{2n}\xi^{2n} + \sum_{n=0}^{\infty} a_{2n+1}\xi^{2n+1}$$

O termo a_{2n} possui dependência da constante a_0 enquanto que o valor de a_{2n+1} depende de a_1. Se $2N+1-K=0$ para um N par, isso significa que, da relação de recorrência (312), a série dos coeficientes pares a_{2n} será truncada, mas a série dos coeficientes ímpares continua tendendo a uma exponencial. Isso pode ser evitado fazendo $a_1 = 0$. O procedimento é análogo se $2N+1-K=0$ para um N ímpar.

A expressão $2N+1-K=0$ implica

$$\begin{aligned} 2N+1 &= K \\ &= \frac{2E_N}{\hbar\omega} \\ E_N &= \hbar\omega\left(N+\frac{1}{2}\right) \end{aligned}$$

ou seja, a energia não pode ter qualquer valor e, portanto, é quantizada.

Quando $N=0$, fazemos $a_1 = 0$ e então

$$\Psi_0(\xi) = \sum_{n=0}^{0} a_n \xi^n e^{-\xi^2/2} = a_0 e^{-\xi^2/2}$$

Para $N=1$, impomos $a_0 = 0$ e

$$\Psi_1(\xi) = \sum_{n=0}^{1} a_n \xi^n e^{-\xi^2/2} = a_1 \xi e^{-\xi^2/2}$$

Por conveniência, vamos multiplicar cada solução particular por 2^N, de forma que

$$\Psi_1(\xi) = 2a_1\xi e^{-\xi^2/2}$$

Para obtermos as demais soluções, basta usar a relação de recorrência para expressar $a_2, a_3, a_4, \ldots$ em termos de a_0 ou a_1:

$$\begin{aligned} \Psi_2(\xi) &= -\frac{a_0}{2}(4\xi^2 - 2)e^{-\xi^2/2} \\ \Psi_3(\xi) &= -\frac{a_1}{12}(8\xi^3 - 12\xi)e^{-\xi^2/2} \\ \Psi_4(\xi) &= \frac{a_0}{12}(16\xi^4 - 48\xi^2 + 12)e^{-\xi^2/2} \\ \Psi_5(\xi) &= \frac{a_1}{120}(32\xi^5 - 160\xi^3 + 120\xi)e^{-\xi^2/2} \end{aligned}$$

Os termos entre parênteses de cada solução particular são os polinômios de Hermite.

A solução geral da equação de Schrödinger independente do tempo do oscilador harmônico quântico é, portanto, a somatória das soluções particulares:

$$\Psi(\xi) = \left[\sum_{n=0}^{\infty} c_n H_n(\xi)\right] e^{-\xi^2/2} \tag{315}$$

onde H_n é o polinômio de Hermite de grau n e os valores c_n são coeficientes arbitrários.

Por outro lado, fazendo $p(\xi) = 1$ no problema de Sturm-Liouville, obtemos

$$\begin{aligned} Lu + \lambda ru &= (pu')' + qu + \lambda ru \\ &= u'' + [q + \lambda r]u \end{aligned}$$

e como $Lu + \lambda ru = 0$, temos a equação diferencial

$$u'' + [q + \lambda r]u = 0$$

Fazendo $r(\xi) = 1$ e $q(\xi) = -\xi^2$, chegamos na expressão (306), que é a EDO do oscilador harmônico quântico

$$u''(\xi) + \left[\lambda - \xi^2\right] u(\xi) = 0 \quad \text{onde} \quad \lambda = K = 2n + 1 \quad n \geq 0 \tag{316}$$

As autofunções são

$$u_n(\xi) = d_n H_n(\xi) e^{-\xi^2/2} \tag{317}$$

onde as constantes d_n precisam ser escolhidas de modo que $\|u_n\| = 1 \quad \forall n$. Para isso, podemos usar o fato de que

$$\int_{-\infty}^{\infty} H_n^2(s) e^{-s^2} \, \mathrm{d}s = 2^n n! \sqrt{\pi} \tag{318}$$

A demonstração da relação (318) está no apêndice (G.1).

Note que

$$\lim_{|x| \to \infty} \sqrt{r(\xi)} u_n(\xi) = \lim_{|x| \to \infty} d_n H_n(\xi) e^{-\xi^2/2} = 0$$

o que significa que podemos usar os resultados da teoria de Sturm-Liouville para essas autofunções u_n.

Observando (318) e sabendo que

$$\|u_n\|^2 = \int_{-\infty}^{\infty} d_n^2 H_n^2(s) e^{-s^2} \, \mathrm{d}s = d_n^2 \left(2^n n! \sqrt{\pi}\right) = 1$$

concluímos que

$$d_n^2 = \frac{1}{2^n n! \sqrt{\pi}}$$

A relação de completude dessas autofunções nos diz que se g é uma função quadrado integrável em todo o $\mathbb{R}$ (isto é, pertence a $\mathcal{L}^2$), então

$$g(\xi) = \sum_{n=0}^{\infty} \langle u_n, g \rangle \, u_n(\xi) \tag{319}$$

onde

$$\langle u_n, g \rangle = \int_{-\infty}^{\infty} u_n(s) g(s) \, \mathrm{d}s$$

Agora vamos considerar por exemplo $g(\xi) = h(\xi)e^{-\xi^2/2}$, onde $h(\xi)$ é uma função que torne $g(\xi)$ quadrado integrável em todo o $\mathbb{R}$, como por exemplo um polinômio ou uma função limitada. Substituindo a autofunção em (319), obtemos

$$\begin{aligned}
h(\xi)e^{-\xi^2/2} &= \sum_{n=0}^{\infty} \langle u_n, h(\xi)e^{-\xi^2/2}\rangle\, d_n H_n(\xi)e^{-\xi^2/2} \\
&= \left[\sum_{n=0}^{\infty}\left(\int_{-\infty}^{\infty} d_n H_n(s)e^{-s^2/2}h(s)e^{-s^2/2}\,\mathrm{d}s\right) d_n H_n(\xi)\right] e^{-\xi^2/2} \\
&= \left[\sum_{n=0}^{\infty}\left(d_n^2\int_{-\infty}^{\infty} h(s)H_n(s)e^{-s^2}\,\mathrm{d}s\right) H_n(\xi)\right] e^{-\xi^2/2} \\
&= \left[\sum_{n=0}^{\infty}\left(\frac{1}{2^n n!\sqrt{\pi}}\int_{-\infty}^{\infty} h(s)H_n(s)e^{-s^2}\,\mathrm{d}s\right) H_n(\xi)\right] e^{-\xi^2/2}
\end{aligned}$$

Comparando a última expressão com (315), em especial o termo entre parênteses, chegamos à conclusão que a escolha

$$c_n = \frac{1}{2^n n!\sqrt{\pi}}\int_{-\infty}^{\infty} h(s)H_n(s)e^{-s^2}\,\mathrm{d}s \tag{320}$$

fará a solução geral $\Psi(\xi)$ em (315) convergir uniformemente para $h(\xi)e^{-s^2/2}$ em todo o $\mathbb{R}$.

Exemplo 5.1. *Para a função $h(\xi) = \xi^3$, temos*

$$c_0 = \frac{1}{2^0 0!\sqrt{\pi}}\int_{-\infty}^{\infty} h(s)H_0(s)e^{-s^2}\,\mathrm{d}s = \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} s^3e^{-s^2}\,\mathrm{d}s = 0$$

$$c_1 = \frac{1}{2^1 1!\sqrt{\pi}}\int_{-\infty}^{\infty} h(s)H_1(s)e^{-s^2}\,\mathrm{d}s = \frac{1}{2\sqrt{\pi}}\int_{-\infty}^{\infty} 2s^4e^{-s^2}\,\mathrm{d}s = \frac{3}{4}$$

$$c_2 = \frac{1}{2^2 2!\sqrt{\pi}}\int_{-\infty}^{\infty} h(s)H_2(s)e^{-s^2}\,\mathrm{d}s = \frac{1}{8\sqrt{\pi}}\int_{-\infty}^{\infty} s^3(4s^2-2)e^{-s^2}\,\mathrm{d}s = 0$$

Como e^{-s^2} é uma função par, $s^3H_n(s)$ é ímpar quando n é par e o intervalo de integração é simétrico no eixo real, segue que todos os coeficientes pares c_{2n} são nulos.

$$c_3 = \frac{1}{2^3 3!\sqrt{\pi}}\int_{-\infty}^{\infty} h(s)H_3(s)e^{-s^2}\,\mathrm{d}s = \frac{1}{48\sqrt{\pi}}\int_{-\infty}^{\infty} s^3(8s^3-12s)e^{-s^2}\,\mathrm{d}s = \frac{1}{8}$$

$$\begin{aligned}
c_5 &= \frac{1}{2^5 5!\sqrt{\pi}}\int_{-\infty}^{\infty} h(s)H_5(s)e^{-s^2}\,\mathrm{d}s \\
&= \frac{1}{2^5 5!\sqrt{\pi}}\int_{-\infty}^{\infty} s^3(32s^5-160s^3+120s)e^{-s^2}\,\mathrm{d}s = 0
\end{aligned}$$

$$c_7 = \frac{1}{2^7 7! \sqrt{\pi}} \int_{-\infty}^{\infty} s^3(128s^7 - 1344s^5 + 3360s^3 - 1680s)e^{-s^2}\,\mathrm{d}s = 0$$

$$c_9 = d_9^2 \int_{-\infty}^{\infty} s^3(512s^9 - 9216s^7 + 48384s^5 - 80640s^3 + 30240s)e^{-s^2}\,\mathrm{d}s = 0$$

Note que todos os coeficientes após c_3 são aparentam ser nulos, e isso é verdade pois

$$\int_{-\infty}^{\infty} s^k H_n(s)e^{-s^2}\,\mathrm{d}s = 0 \qquad se \qquad k < n \tag{321}$$

A demonstração de (321) *está no apêndice* (G.2).

Finalmente, temos que

$$\begin{aligned}
\Psi(\xi) &= \left[\sum_{n=0}^{\infty} c_n H_n(\xi)\right] e^{-\xi^2/2} \\
&= \left[c_1 H_1(\xi) + c_3 H_3(\xi)\right] e^{-\xi^2/2} \\
&= \left[\frac{3}{4}(2\xi) + \frac{1}{8}(8\xi^3 - 12\xi)\right] e^{-\xi^2/2} \\
&= \left[\frac{6}{4}\xi + \frac{1}{8}(8\xi^3) - \frac{6}{4}\xi\right] e^{-\xi^2/2} \\
&= \xi^3 e^{-\xi^2/2}
\end{aligned}$$

Note que o resultado coincidiu totalmente com a função $h(\xi)e^{-\xi^2/2}$ uma vez que $h(\xi) = \xi^3$ pertence ao espaço gerado pelos polinômios de Hermite (que no caso é o espaço de todos os polinômios). O gráfico de $\Psi(\xi)$ está na figura (7).

Sabendo que

$$\xi = \sqrt{\frac{m\omega}{\hbar}}(x - x_0)$$

então a solução da equação diferencial

$$\Psi''(x - x_0) + \left[\frac{2mE}{\hbar^2} - \frac{m^2}{\hbar^2}\omega^2(x - x_0)^2\right]\Psi(x - x_0) = 0 \tag{322}$$

é dada por

$$\Psi(x) = \left(\sqrt{\frac{m\omega}{\hbar}}\right)^3 (x - x_0)^3 \exp\left(-\frac{1}{2}\left(\sqrt{\frac{m\omega}{\hbar}}\right)^2 (x - x_0)^2\right)$$

Para que $\Psi(x)$ se torne uma solução física (isto é, descreva o comportamento de uma partícula) é necessário normalizá-la multiplicando por alguma constante de forma que $\Psi(x)$ satisfaça

$$\int_{-\infty}^{\infty} |\Psi(x)|^2\,\mathrm{d}x = 1$$
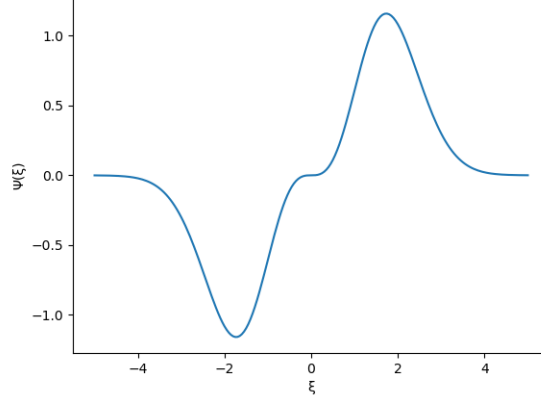
Figura 7: A solução da equação diferencial do oscilador harmônico quântico (não-normalizado) $\xi^3 e^{-\xi^2/2}$ escrita como uma combinação linear de polinômios de Hermite com exponencial.

A função de onda $\psi(x,t)$ pode ser obtida ao multiplicar $\Psi(x)$ por $\exp\left(-i\frac{E}{\hbar}t\right)$.

Exemplo 5.2. *Para a função* $h(\xi) = \cos(\xi)$, *temos*

$$c_0 = \frac{1}{2^0 0!\sqrt{\pi}} \int_{-\infty}^{\infty} h(s)H_0(s)e^{-s^2}\,\mathrm{d}s = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \cos(s)e^{-s^2}\,\mathrm{d}s = \frac{1}{\sqrt[4]{e}}$$

$$c_1 = \frac{1}{2^1 1!\sqrt{\pi}} \int_{-\infty}^{\infty} h(s)H_1(s)e^{-s^2}\,\mathrm{d}s = \frac{1}{2\sqrt{\pi}} \int_{-\infty}^{\infty} \cos(s)2se^{-s^2}\,\mathrm{d}s = 0$$

$$c_2 = \frac{1}{2^2 2!\sqrt{\pi}} \int_{-\infty}^{\infty} h(s)H_2(s)e^{-s^2}\,\mathrm{d}s = \frac{1}{8\sqrt{\pi}} \int_{-\infty}^{\infty} \cos(s)(4s^2-2)e^{-s^2}\,\mathrm{d}s = -\frac{1}{8\sqrt[4]{e}}$$

$$c_3 = \frac{1}{2^3 3!\sqrt{\pi}} \int_{-\infty}^{\infty} h(s)H_3(s)e^{-s^2}\,\mathrm{d}s = \frac{1}{48\sqrt{\pi}} \int_{-\infty}^{\infty} \cos(s)(8s^3 - 12s)e^{-s^2}\,\mathrm{d}s = 0$$

Como e^{-s^2} *é uma função par,* $\cos(s)H_n(s)$ *é ímpar quando* n *é ímpar e o intervalo de integração é simétrico no eixo real, segue que todos os coeficientes ímpares* c_{2n+1} *são nulos.*

$$c_4 = \frac{1}{2^4 4!\sqrt{\pi}} \int_{-\infty}^{\infty} \cos(s)(16s^4 - 48s^2 + 12)e^{-s^2}\,\mathrm{d}s = \frac{1}{384\sqrt[4]{e}}$$

$$c_6 = \frac{1}{2^6 6!\sqrt{\pi}} \int_{-\infty}^{\infty} \cos(s)(64s^6 - 480s^4 + 720s^2 - 120)e^{-s^2}\,\mathrm{d}s = -\frac{1}{46080\sqrt[4]{e}}$$

Note que os valores das integrais valem $\frac{\sqrt{\pi}}{\sqrt[4]{e}}$ *ou* $-\frac{\sqrt{\pi}}{\sqrt[4]{e}}$ *e aparentam oscilar entre eles a cada coeficiente par. Para fins práticos, somente esses seis primeiros*

coeficientes são suficientes para se aproximar da curva $\cos(\xi)e^{-\xi^2/2}$ *conforme pode ser visto nas figuras* (8) *e* (9).

A solução final é dada por

$$\begin{aligned}
\Psi(\xi) &= \left[\sum_{n=0}^{\infty} c_n H_n(\xi)\right] e^{-\xi^2/2} \\
&= \left[\sum_{n=0}^{\infty} c_{2n} H_{2n}(\xi)\right] e^{-\xi^2/2} \\
&= \left[c_0 H_0(\xi) + c_2 H_2(\xi) + c_4 H_4(\xi) + c_6 H_6(\xi) + \ldots\right] e^{-\xi^2/2}
\end{aligned}$$

$$\Psi(\xi) = \left[\frac{1}{\sqrt[4]{e}} - \frac{1}{8\sqrt[4]{e}}(4\xi^2 - 2) + \frac{1}{384\sqrt[4]{e}}(16\xi^4 - 48\xi^2 + 12) - \frac{1}{46080\sqrt[4]{e}}(64\xi^6 - 480\xi^4 + 720\xi^2 - 120) + \ldots\right] e^{-\xi^2/2} \quad (323)$$

A solução da equação diferencial (322) *é obtida ao substituir a variável adimensional*

$$\xi = \sqrt{\frac{m\omega}{\hbar}}(x - x_0)$$

em (323). *E para tornar a solução matemática uma solução física, é necessário normalização.*
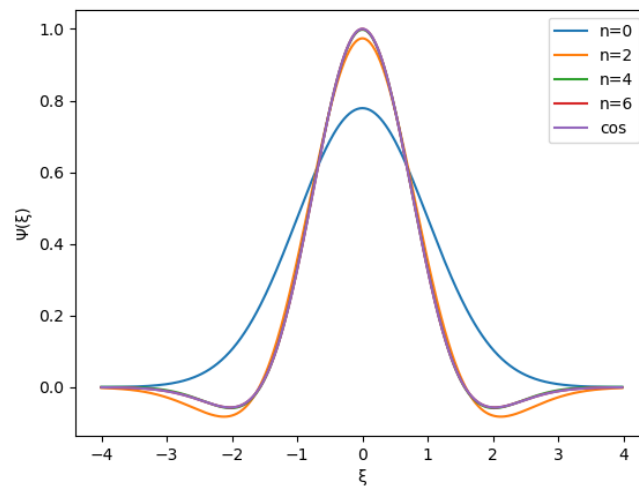
Figura 8: A solução da equação diferencial do oscilador harmônico quântico (não-normalizado) convergindo uniformemente para $\cos(\xi)e^{-\xi^2/2}$ como uma combinação linear de polinômios de Hermite (até o grau 6) com exponencial. A curva roxa é a função para onde $\Psi(\xi)$ está convergindo uniformemente, ou seja, $\cos(\xi)e^{-\xi^2/2}$ enquanto as outras são as combinações lineares em convergência.
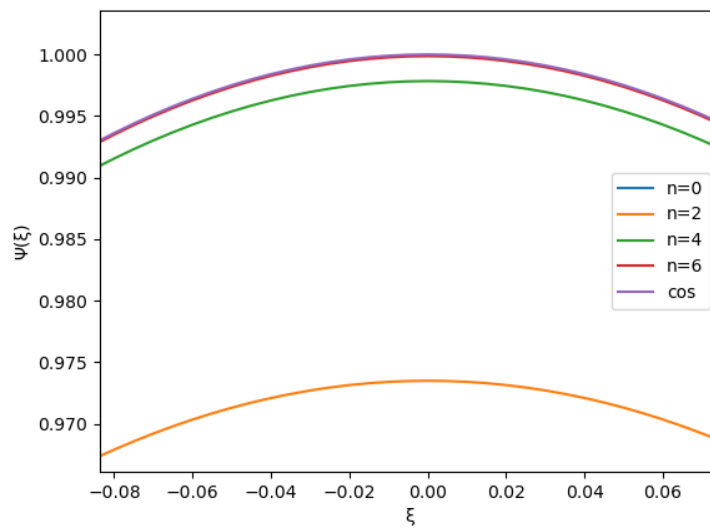
Figura 9: Visão mais detalhada da convergência uniforme para $\cos(\xi)e^{-\xi^2/2}$ da combinação linear de polinômios de Hermite (até o grau 6) com exponencial. A linha vermelha, correspondente ao grau 6, quase coincide com a linha roxa.

6 Partícula confinada em um vácuo esférico

Nesta seção, iremos introduzir os harmônicos esféricos, funções que aparecem frequentemente em soluções de equações diferenciais de problemas da física.

Um desses problemas é a de uma partícula confinada em um vácuo esférico, que pode ser descrito pelo potencial

$$V(r) = \begin{cases} 0 & , \text{se } 0 \leq r \leq r_0 \\ +\infty & r_0 > a \end{cases}$$

Na região dentro da esfera, a equação independente do tempo de Schrödinger retorna

$$-\frac{\hbar^2}{2m_0}\nabla^2\psi - E\psi = 0$$

e portanto a equação diferencial que devemos resolver é

$$\nabla^2\psi + \frac{2m_0E}{\hbar^2}\psi = 0 \tag{324}$$

onde

$$\nabla^2\psi = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right) + \frac{1}{r^2\sin(\theta)}\frac{\partial}{\partial\theta}\left(\sin(\theta)\frac{\partial\psi}{\partial\theta}\right) + \frac{1}{r^2\sin^2(\theta)}\frac{\partial^2\psi}{\partial\varphi^2}$$

O objetivo é descobrir quais são os níveis de energia permitidos pelo sistema físico. Como $\min\{V(r)\} = 0$, o teorema (3.4) garante que os níveis de energia precisam ser positivos para que ψ seja normalizável.

6.1 Separação das variáveis esféricas

Abrindo as expressões em parênteses da equação (324), obtemos

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right) = \frac{1}{r^2}\left[2r\frac{\partial\psi}{\partial r} + r^2\frac{\partial^2\psi}{\partial r^2}\right] = \frac{2}{r}\frac{\partial\psi}{\partial r} + \frac{\partial^2\psi}{\partial r^2}$$

$$\begin{aligned}\frac{1}{r^2\sin(\theta)}\frac{\partial}{\partial\theta}\left(\sin(\theta)\frac{\partial\psi}{\partial\theta}\right) &= \frac{1}{r^2\sin(\theta)}\left[\cos(\theta)\frac{\partial\psi}{\partial\theta} + \sin(\theta)\frac{\partial^2\psi}{\partial\theta^2}\right] \\ &= \frac{\cot(\theta)}{r^2}\frac{\partial\psi}{\partial\theta} + \frac{1}{r^2}\frac{\partial^2\psi}{\partial\theta^2}\end{aligned}$$

Portanto a equação diferencial se torna

$$\frac{2}{r}\frac{\partial\psi}{\partial r} + \frac{\partial^2\psi}{\partial r^2} + \frac{\cot(\theta)}{r^2}\frac{\partial\psi}{\partial\theta} + \frac{1}{r^2}\frac{\partial^2\psi}{\partial\theta^2} + \frac{1}{r^2\sin^2(\theta)}\frac{\partial^2\psi}{\partial\varphi^2} + \frac{2m_0E}{\hbar^2}\psi = 0 \tag{325}$$

$$. \tag{326}$$

Aplicando separação de variáveis na equação diferencial (325), com $\psi(r,\theta,\varphi) = R(r)\Theta(\theta)\Phi(\varphi)$, obtemos

$$\frac{2}{r}R'(r)\Theta(\theta)\Phi(\varphi) + R''(r)\Theta(\theta)\Phi(\varphi) + \frac{\cot(\theta)}{r^2}R(r)\Theta'(\theta)\Phi(\varphi) +$$
$$\frac{1}{r^2}R(r)\Theta''(\theta)\Phi(\varphi) + \frac{1}{r^2\sin^2(\theta)}R(r)\Theta(\theta)\Phi''(\varphi) + \frac{2m_0E}{\hbar^2}R(r)\Theta(\theta)\Phi(\varphi) = 0$$

Considerando que $u(r,\theta,\varphi) \neq 0$ implica $R(r) \neq 0$, $\Theta(\theta) \neq 0$ e $\Phi(\varphi) \neq 0$, podemos dividir ambos os lados por $R(r)\Theta(\theta)\Phi(\varphi)$ para chegarmos na expressão

$$\frac{2}{r}\frac{R'(r)}{R(r)} + \frac{R''(r)}{R(r)} + \frac{\cot(\theta)}{r^2}\frac{\Theta'(\theta)}{\Theta(\theta)} + \frac{1}{r^2}\frac{\Theta''(\theta)}{\Theta(\theta)} + \frac{1}{r^2\sin^2(\theta)}\frac{\Phi''(\varphi)}{\Phi(\varphi)} + \frac{2m_0E}{\hbar^2} = 0$$

Isolando primeiro os termos com a variável φ:

$$\frac{2}{r}\frac{R'(r)}{R(r)} + \frac{R''(r)}{R(r)} + \frac{\cot(\theta)}{r^2}\frac{\Theta'(\theta)}{\Theta(\theta)} + \frac{1}{r^2}\frac{\Theta''(\theta)}{\Theta(\theta)} + \frac{2m_0E}{\hbar^2} = -\frac{1}{r^2\sin^2(\theta)}\frac{\Phi''(\varphi)}{\Phi(\varphi)}$$

Multiplicando ambos os lados por $r^2\sin^2(\theta)$,

$$(r^2\sin^2(\theta))\cdot\left(\frac{2}{r}\frac{R'(r)}{R(r)} + \frac{R''(r)}{R(r)} + \frac{\cot(\theta)}{r^2}\frac{\Theta'(\theta)}{\Theta(\theta)} + \frac{1}{r^2}\frac{\Theta''(\theta)}{\Theta(\theta)} + \frac{2m_0E}{\hbar^2}\right) = -\frac{\Phi''(\varphi)}{\Phi(\varphi)}$$

Como o primeiro membro só possui dependência em r e θ enquanto que o segundo membro só depende de φ, segue que cada membro é igual a uma mesma constante λ_1:

$$(r^2\sin^2(\theta))\cdot\left(\frac{2}{r}\frac{R'(r)}{R(r)} + \frac{R''(r)}{R(r)} + \frac{\cot(\theta)}{r^2}\frac{\Theta'(\theta)}{\Theta(\theta)} + \frac{1}{r^2}\frac{\Theta''(\theta)}{\Theta(\theta)} + \frac{2m_0E}{\hbar^2}\right) = \lambda_1 \tag{327}$$

e

$$-\frac{\Phi''(\varphi)}{\Phi(\varphi)} = \lambda_1$$

A última expressão é a equação diferencial

$$\Phi''(\varphi) = -\lambda_1\Phi(\varphi)$$

Dividindo ambos os lados por $\sin^2(\theta)$ em (327), obtemos

$$2r\frac{R'(r)}{R(r)} + r^2\frac{R''(r)}{R(r)} + \cot(\theta)\frac{\Theta'(\theta)}{\Theta(\theta)} + \frac{\Theta''(\theta)}{\Theta(\theta)} + \frac{2m_0E}{\hbar^2}r^2 = \frac{\lambda_1}{\sin^2(\theta)}$$

Isolando os termos com θ e r para cada membro, temos que

$$2r\frac{R'(r)}{R(r)} + r^2\frac{R''(r)}{R(r)} + \frac{2m_0E}{\hbar^2}r^2 = -\left[\frac{\Theta''(\theta)}{\Theta(\theta)} + \cot(\theta)\frac{\Theta'(\theta)}{\Theta(\theta)} - \frac{\lambda_2}{\sin^2(\theta)}\right]$$

ou seja, cada membro é uma constante λ_2:

$$\frac{\Theta''(\theta)}{\Theta(\theta)} + \cot(\theta)\frac{\Theta'(\theta)}{\Theta(\theta)} - \frac{\lambda_1}{\sin^2(\theta)} = -\lambda_2 \tag{328}$$

$$2r\frac{R'(r)}{R(r)} + r^2\frac{R''(r)}{R(r)} + \frac{2mE}{\hbar^2}r^2 = \lambda_2 \tag{329}$$

Na expressão (328), vamos multiplicar ambos os lados por $\Theta(\theta)$ e então obteremos a equação diferencial

$$\Theta''(\theta) + \cot(\theta)\Theta'(\theta) + \left(\lambda_2 - \frac{\lambda_1}{\sin^2(\theta)}\right)\Theta(\theta) = 0$$

E da equação (329) obtemos a última EDO ao multiplicar por $R(r)$:

$$r^2R''(r) + 2rR'(r) + \left[\frac{2m_0E}{\hbar^2}r^2 - \lambda_2\right]R(r) = 0$$

Com isso, temos três equações diferenciais ordinárias:

$$\Phi''(\varphi) = -\lambda_1\Phi(\varphi) \tag{330}$$

$$\Theta''(\theta) + \cot(\theta)\Theta'(\theta) + \left(\lambda_2 - \frac{\lambda_1}{\sin^2(\theta)}\right)\Theta(\theta) = 0 \tag{331}$$

$$r^2R''(r) + 2rR'(r) + \left(\frac{2m_0E}{\hbar^2}r^2 - \lambda_2\right)R(r) = 0 \tag{332}$$

6.2 Equação azimutal

Para a variável φ, temos a EDO

$$\Phi''(\varphi) = -\lambda_1\Phi(\varphi) \tag{333}$$

Fazendo $\Phi(\varphi) = e^{i\gamma\varphi}$ e encontrando os valores de γ, encontramos a solução geral

$$\Phi(\varphi) = A\cdot e^{i\gamma_1\cdot\varphi} + B\cdot e^{i\gamma_2\cdot\varphi}$$

com $\gamma_1 = \sqrt{\lambda_1}$ e $\gamma_2 = -\sqrt{\lambda_1}$. O fator imaginário i aparece nas exponenciais devido ao sinal negativo de (333).

Como φ denota o ângulo azimutal na função de onda, segue que é preciso impor $\Phi(\varphi + 2\pi) = \Phi(\varphi)\ \forall\phi \in [0, 2\pi]$ como condição de fronteira para que a solução seja contínua.

Considerando primeiramente $\lambda_1 < 0$, temos $\gamma_1 = i\sqrt{|\lambda_1|}$ e $\gamma_2 = -i\sqrt{|\lambda_1|}$, no que resulta

$$\Phi(\varphi) = A\cdot e^{-\sqrt{|\lambda_1|}\cdot\varphi} + B\cdot e^{\sqrt{|\lambda_1|}\cdot\varphi}$$

Impondo a condição $\Phi(\varphi + 2\pi) = \Phi(\varphi)\ \forall\phi \in [0, 2\pi]$, obtemos

$$A \cdot \exp\Big\{-\sqrt{|\lambda_1|} \cdot \varphi - 2\pi\sqrt{|\lambda_1|}\Big\} + B \cdot \exp\Big\{\sqrt{|\lambda_1|} \cdot \varphi + 2\pi\sqrt{|\lambda_1|}\Big\} = A \cdot \exp\Big\{-\sqrt{|\lambda_1|} \cdot \varphi\Big\} + B \cdot \exp\Big\{\sqrt{|\lambda_1|} \cdot \varphi\Big\}$$

Dos termos proporcionais a $\exp\Big\{\sqrt{|\lambda_1|} \cdot \varphi\Big\}$, temos

$$B \cdot e^{2\pi\sqrt{|\lambda_1|}} = B \implies e^{2\pi\sqrt{|\lambda_1|}} = 1$$

Mas isso só é possível se $\lambda_1 = 0$, o que não convém ao caso e então $\lambda_1 < 0$ não resulta em soluções não-nulas.

Se $\lambda_1 = 0$, temos o termo constante $\Phi(\varphi) = C$, que trivialmente obedece às condições de contorno.

Considerando agora $\lambda_1 > 0$, temos $\sqrt{\lambda_1} \in \mathbb{R}$ e

$$\Phi(\varphi) = A \cdot e^{i\sqrt{\lambda_1}\cdot\varphi} + B \cdot e^{-i\sqrt{\lambda_1}\cdot\varphi}$$

Analogamente ao caso $\lambda_1 < 0$, a condição de contorno resulta na equação

$$e^{2\pi i\sqrt{\lambda_1}} = 1$$

Da identidade de Euler, temos que

$$\cos\Big(2\pi\sqrt{\lambda_1}\Big) + i\sin\Big(2\pi\sqrt{\lambda_1}\Big) = 1 \tag{334}$$

Note que no primeiro membro aparece um termo real e outro imaginário enquanto que o segundo membro é um número real. Isso significa que

$$\sin\Big(2\pi\sqrt{\lambda_1}\Big) = 0 \implies 2\pi\sqrt{\lambda_1} = 2\pi m \quad m \in \mathbb{Z}$$

e portanto

$$\lambda_{1_m} = m^2$$

Note que esse λ_{1_m} também satisfaz a parte real da equação (334) e que dentre os valores de m inclui o zero, que já aborda o caso $\Phi(\varphi) = C$.

A solução geral da EDO (330) que diz respeito a φ é, portanto

$$\Phi_m(\varphi) = A \cdot e^{im\varphi} + B \cdot e^{-im\varphi} \quad \forall m \in \mathbb{Z} \tag{335}$$

Se m for negativo, obteremos a mesma solução para m positivo, pois as duas soluções independentes são o conjugado da outra. Então podemos simplificar o resultado para

$$\Phi_m(\varphi) = A \cdot e^{im\varphi} \quad \forall m \in \mathbb{Z}$$

Se em algum problema essas funções representam algo real, como um potencial elétrico, precisaríamos transformar essas exponenciais imaginárias em senos e cossenos reais usando $\overline{A} = B$ na equação (335). Mas como aqui a intenção é descrever uma função de onda ψ, iremos deixar a exponencial imaginária por conveniência.

6.3 Equação de Legendre associada

Sabendo que $\lambda_1 = m^2$, com m inteiro, a equação diferencial para θ (331) se torna

$$\Theta''(\theta) + \cot(\theta)\Theta'(\theta) + \left(\lambda_2 - \frac{m^2}{\sin^2(\theta)}\right)\Theta(\theta) = 0 \tag{336}$$

Dividindo ambos os lados da equação por $\Theta(\theta)$, obtemos

$$\frac{1}{\Theta(\theta)}\left[\Theta''(\theta) + \cot(\theta)\Theta'(\theta)\right] + \lambda_2 - \frac{m^2}{\sin^2(\theta)} = 0 \tag{337}$$

Agora note que

$$\begin{aligned}\Theta''(\theta) + \cot(\theta)\Theta'(\theta) &= \frac{1}{\sin(\theta)}\left[\sin(\theta)\Theta''(\theta) + \cos(\theta)\Theta'(\theta)\right] \\ &= \frac{1}{\sin(\theta)}\frac{d}{d\theta}\left(\sin(\theta)\Theta'(\theta)\right)\end{aligned}$$

Substituindo a última expressão em (337), obtemos

$$\frac{1}{\sin(\theta)\Theta(\theta)}\left[\frac{d}{d\theta}\left(\sin(\theta)\Theta'(\theta)\right)\right] + \lambda_2 - \frac{m^2}{\sin^2(\theta)} = 0$$

Fazendo a substituição $x = \cos(\theta)$, temos que

$$\sin^2(\theta) = 1 - \cos^2(\theta) = 1 - x^2$$

$$\frac{dx}{d\theta} = \sin(\theta)$$

$$\frac{d}{d\theta} = \frac{d}{dx}\frac{dx}{d\theta} = \sin(\theta)\frac{d}{dx}$$

$$\frac{d\Theta}{d\theta} = \frac{d\Theta}{dx}\frac{dx}{d\theta} = \frac{d\Theta}{dx}\sin(\theta) = \sin(\theta)\Theta'(x)$$

e a equação diferencial fica

$$\frac{1}{\sin(\theta)\Theta(x)}\left[\sin(\theta)\frac{d}{dx}\left((1-x^2)\Theta'(x)\right)\right] + \lambda_2 - \frac{m^2}{1-x^2} = 0$$

$$\frac{1}{\Theta(x)}\frac{d}{dx}\left((1-x^2)\Theta'(x)\right) + \lambda_2 - \frac{m^2}{1-x^2} = 0$$

Multiplicando por $\Theta(x)$, chegamos na expressão

$$\frac{d}{dx}\left((1-x^2)\Theta'(x)\right) + \left[\lambda_2 - \frac{m^2}{1-x^2}\right]\Theta(x) = 0 \tag{338}$$

As soluções da equação diferencial (338) são denominadas de funções associadas de Legendre.

Na forma

$$(1-x^2)\Theta''(x) - 2x\Theta'(x) + \left[\lambda_2 - \frac{m^2}{1-x^2}\right]\Theta(x) = 0 \tag{339}$$

podemos ver que a equação diferencial possui singularidades em $x = \pm 1$. Então iremos precisar do seguinte teorema:

Teorema 6.1 (Caso inteiro do Teorema de Frobenius)**.** *Sejam $P(x)$ e $Q(x)$ funções tais que*

- $\lim_{x\to x_0}(x-x_0)P(x) = p_0 \in \mathbb{R}$
- $\lim_{x\to x_0}(x-x_0)^2Q(x) = q_0 \in \mathbb{R}$

e que $(x-x_0)P(x)$ e $(x-x_0)^2Q(x)$ sejam funções analíticas no intervalo $|x-x_0| < R$, com x_0 sendo um ponto singular regular.

Se as raízes α_1 e α_2 da equação indicial $\alpha(\alpha-1)+p_0\alpha+q_0 = 0$ são tais que $\alpha_1 - \alpha_2 = C$, com C inteiro não-negativo, então a equação diferencial

$$f''(x) + P(x)f'(x) + Q(x)f(x) = 0$$

possui soluções duas soluções linearmente independentes $f_1(x)$ e $f_2(x)$ dadas por

$$f_1(x) = |x-x_0|^{\alpha_1}\sum_{n=0}^{\infty} a_n(x-x_0)^n$$

e

$$f_2(x) = |x-x_0|^{\alpha_2}\sum_{n=0}^{\infty} b_n(x-x_0)^n + Kf_1(x)\ln|x-x_0|$$

onde K é uma constante e é igual a 1 se $\alpha_1-\alpha_2 = 0$. Ambas convergem em $|x-x_0| < R$.

Na forma

$$\Theta''(x) - \frac{2x}{1-x^2}\Theta'(x) + \left[\frac{\lambda_2}{1-x^2} - \frac{m^2}{(1-x^2)^2}\right]\Theta(x) = 0$$

Podemos fazer

$$P(x) = -\frac{2x}{1-x^2} \quad \text{e} \quad Q(x) = \frac{\lambda_2}{1-x^2} - \frac{m^2}{(1-x^2)^2}$$

para vermos se $x_0 = 1$ e $x_1 = -1$ são singulares regulares nesses pontos. Sabendo que $1-x^2 = (1+x)(1-x)$, temos

$$(x-1)P(x) = -(x-1)\cdot\frac{2x}{(1-x)(1+x)} = \frac{2x}{1+x}$$

que é analítico no intervalo em que $|x| < 1$, ou seja, em $(-1, 1)$. E então

$$p_0 = \lim_{x \to 1} (x-1)P(x) = \lim_{x \to 1} \frac{2x}{1+x} = 1 \in \mathbb{R}$$

Para $Q(x)$, temos

$$\begin{aligned}(x-1)^2 Q(x) &= \frac{(x-1)(x-1)\lambda_2}{(1-x)(1+x)} - \frac{(x-1)(x-1)m^2}{(1-x)(1-x)(1+x)(1+x)} \\ &= -\frac{(x-1)\lambda_2}{(1+x)} - \frac{m^2}{(1+x)(1+x)}\end{aligned}$$

que é analítico no intervalo $(-1, 1)$. Portanto

$$q_0 = \lim_{x \to 1} (x-1)^2 Q(x) = \lim_{x \to 1} -\frac{(x-1)\lambda_2}{(1+x)} - \frac{m^2}{(1+x)(1+x)} = -\frac{m^2}{4} \in \mathbb{R}$$

Então $x_0 = 1$ é um ponto singular regular, com $p_0 = 1$ e $q_0 = -\left(\frac{m}{2}\right)^2$. Pondo esses valores na equação indicial, temos que

$$\alpha(\alpha - 1) + \alpha - \left(\frac{m}{2}\right)^2 = 0$$

$$\alpha^2 = \left(\frac{m}{2}\right)^2$$

no que implica $\alpha_1 = \frac{m}{2}$ e $\alpha_2 = -\frac{m}{2}$.

Analogamente, para $x_1 = -1$, podemos ver que

$$p_1 = 1 \in \mathbb{R} \quad \text{e} \quad q_1 = -\frac{m^2}{4} \in \mathbb{R}$$

que vão gerar as mesmas soluções para α.

Como $\alpha_1 - \alpha_2 = m > 0$ e m é um inteiro, devemos usar o caso inteiro do teorema de Frobenius. Vamos supor inicialmente que m é não-negativo, pois assim os fatores $|x-1|^{m/2}$ ou $|x+1|^{m/2}$ não divergirão em $x = \pm 1$. Então a primeira solução da equação diferencial (338) para $x_0 = 1$ possui uma forma do tipo

$$f_{1_1}(x) = |x-1|^{m/2} \sum_{n=0}^{\infty} e_n (x-1)^n \tag{340}$$

e para $x_1 = -1$,

$$f_{1_{-1}}(x) = |x+1|^{m/2} \sum_{n=0}^{\infty} g_n (x+1)^n \tag{341}$$

Sabemos que ambas as soluções são válidas no intervalo $(-1, 1)$. Então segue que nesse intervalo

$$|x-1| = 1 - x \quad \text{e} \quad |x+1| = x + 1$$

no que implica

$$f_{1_1}(x) = (1-x)^{m/2} \sum_{n=0}^{\infty} e_n (x-1)^n \tag{342}$$

e

$$f_{1_{-1}}(x) = (1+x)^{m/2} \sum_{n=0}^{\infty} g_n (x+1)^n \tag{343}$$

Queremos que ambas as funções sejam idênticas no intervalo $(-1,1)$. Sabemos que tanto $(x-1)^n$ quanto $(x+1)^n$ são bases que geram o mesmo espaço gerado por x^n. Então supondo que ambas as séries convergem, podemos rearranjar os termos de forma que

$$f_{1_1}(x) = (1-x)^{m/2} \sum_{n=0}^{\infty} u_n x^n$$

e

$$f_{1_{-1}}(x) = (1+x)^{m/2} \sum_{n=0}^{\infty} v_n x^n$$

para alguns coeficientes u_n e v_n. Isso significa que a solução no intervalo $(-1,1)$ tem tanto uma parcela tanto de $(1-x)^{m/2}$ quanto de $(1+x)^{m/2}$ imbuídas nela. Isso nos sugere tentar

$$f_1(x) = (1-x)^{m/2}(1+x)^{m/2} \sum_{n=0}^{\infty} a_n x^n = (1-x^2)^{m/2} \sum_{n=0}^{\infty} a_n x^n$$

onde a_n são coeficientes a determinar. Vamos substituir $p = \frac{m}{2}$ por conveniência, resultando

$$f_1(x) = (1-x^2)^p \sum_{n=0}^{\infty} a_n x^n$$

Suas derivadas são

$$f_1'(x) = -2xp(1-x^2)^{p-1} \sum_{n=0}^{\infty} a_n x^n + (1-x^2)^p \sum_{n=0}^{\infty} n a_n x^{n-1}$$

$$f_1''(x) = \left(-2xp(1-x^2)^{p-1}\right)' \sum_{n=0}^{\infty} a_n x^n - 2xp(1-x^2)^{p-1} \sum_{n=0}^{\infty} n a_n x^{n-1} - 2xp(1-x^2)^{p-1} \sum_{n=0}^{\infty} n a_n x^{n-1} + (1-x^2)^p \sum_{n=0}^{\infty} n(n-1) a_n x^{n-2}$$

$$f_1''(x) = -2p(1-x^2)^{p-1} \sum_{n=0}^{\infty} a_n x^n + 4x^2 p(p-1)(1-x^2)^{p-2} \sum_{n=0}^{\infty} a_n x^n - 4xp(1-x^2)^{p-1} \sum_{n=0}^{\infty} n a_n x^{n-1} + (1-x^2)^p \sum_{n=0}^{\infty} n(n-1) a_n x^{n-2}$$

Inserindo essas expressões na equação diferencial

$$(1-x^2)f_1''(x) - 2xf_1'(x) + \lambda_2 f(x) - \frac{m^2}{1-x^2}f_1(x) = 0$$

obteremos

$$\begin{aligned}
&-2p(1-x^2)^p\sum_{n=0}^{\infty}a_nx^n + 4x^2p(p-1)(1-x^2)^{p-1}\sum_{n=0}^{\infty}a_nx^n - \\
&\quad 4xp(1-x^2)^p\sum_{n=0}^{\infty}na_nx^{n-1} + (1-x^2)^{p+1}\sum_{n=0}^{\infty}n(n-1)a_nx^{n-2} + \\
&\quad\quad 4x^2p(1-x^2)^{p-1}\sum_{n=0}^{\infty}a_nx^n - 2x(1-x^2)^p\sum_{n=0}^{\infty}na_nx^{n-1} + \\
&\quad\quad\quad \lambda_2(1-x^2)^p\sum_{n=0}^{\infty}a_nx^n - m^2(1-x^2)^{p-1}\sum_{n=0}^{\infty}a_nx^n = 0
\end{aligned}$$

Dividindo ambos os lados por $(1-x^2)^{p-1}$,

$$\begin{aligned}
&-2p(1-x^2)\sum_{n=0}^{\infty}a_nx^n + 4x^2p(p-1)\sum_{n=0}^{\infty}a_nx^n - \\
&\quad 4xp(1-x^2)\sum_{n=0}^{\infty}na_nx^{n-1} + (1-x^2)^2\sum_{n=0}^{\infty}n(n-1)a_nx^{n-2} + \\
&\quad\quad 4x^2p\sum_{n=0}^{\infty}a_nx^n - 2x(1-x^2)\sum_{n=0}^{\infty}na_nx^{n-1} + \\
&\quad\quad\quad \lambda_2(1-x^2)\sum_{n=0}^{\infty}a_nx^n - m^2\sum_{n=0}^{\infty}a_nx^n = 0
\end{aligned}$$

Agrupando os somatórios com a mesma potência,

$$\begin{aligned}
&\Big[-2p(1-x^2) + 4x^2p(p-1) + 4x^2p + \lambda_2(1-x^2) - m^2\Big]\sum_{n=0}^{\infty}a_nx^n + \\
&\quad \Big[-4xp(1-x^2) - 2x(1-x^2)\Big]\sum_{n=0}^{\infty}na_nx^{n-1} + \\
&\quad\quad (1-x^2)^2\sum_{n=0}^{\infty}n(n-1)a_nx^{n-2} = 0
\end{aligned}$$

Simplificando os termos entre parênteses,

$$\left[(2p(2p+1)-\lambda_2)x^2+(\lambda_2-m^2-2p)\right]\sum_{n=0}^{\infty}a_nx^n+$$
$$\left[(4p+2)x^3-(4p+2)x\right]\sum_{n=0}^{\infty}na_nx^{n-1}+$$
$$(x^4-2x^2+1)\sum_{n=0}^{\infty}n(n-1)a_nx^{n-2}=0$$

Distribuindo esses produtos, temos que

$$\sum_{n=0}^{\infty}(2p(2p+1)-\lambda_2)a_nx^{n+2}+\sum_{n=0}^{\infty}(\lambda_2-m^2-2p)a_nx^n+$$
$$\sum_{n=0}^{\infty}(4p+2)na_nx^{n+2}-\sum_{n=0}^{\infty}(4p+2)na_nx^n+\sum_{n=0}^{\infty}n(n-1)a_nx^{n+2}-$$
$$\sum_{n=0}^{\infty}2n(n-1)a_nx^n+\sum_{n=0}^{\infty}n(n-1)a_nx^{n-2}=0$$

Novamente reagrupando os somatórios com a mesma potência,

$$\sum_{n=0}^{\infty}n(n-1)a_nx^{n-2}+$$
$$\sum_{n=0}^{\infty}\left[(\lambda_2-m^2-2p)-(4p+2)n-2n(n-1)\right]a_nx^n+$$
$$\sum_{n=0}^{\infty}\left[2p(2p+1)-\lambda_2+(4p+2)n+n(n-1)\right]a_nx^{n+2}=0$$

Simplificando novamente os termos entre parênteses,

$$\sum_{n=0}^{\infty}n(n-1)a_nx^{n-2}+$$
$$\sum_{n=0}^{\infty}\left[(\lambda_2-m^2)-2n(n+2p)-2p\right]a_nx^n+$$
$$\sum_{n=0}^{\infty}\left[2p(2p+1)-\lambda_2+n[n+(4p+1)]\right]a_nx^{n+2}=0$$

Retirando os quatro primeiros termos do primeiro somatório e os dois primeiros

do segundo somatório, temos

$$0+0+2a_2+6a_3x+[(\lambda_2-m^2)-2p]a_0+[(\lambda_2-m^2)-2-6p]a_1x+$$
$$\sum_{n=4}^{\infty} n(n-1)a_n x^{n-2}+$$
$$\sum_{n=2}^{\infty}\Big[(\lambda_2-m^2)-2n(n+2p)-2p\Big]a_n x^n+$$
$$\sum_{n=0}^{\infty}\Big[2p(2p+1)-\lambda_2+n[n+(4p+1)]\Big]a_n x^{n+2}=0$$

Recolocando a constante $m=2p$ e descendo os índices dos dois primeiros somatórios,

$$\Big[[(\lambda_2-m^2)-m]a_0+2a_2\Big]x^0+\Big[6a_3+[(\lambda_2-m^2)-2-3m]a_1\Big]x^1+$$
$$\sum_{n=0}^{\infty}\Big[(n+4)(n+3)a_{n+4}+[(\lambda_2-m^2)-2(n+2)(n+m+2)-m]a_{n+2}+$$
$$[m(m+1)-\lambda_2+n(n+2m+1)]a_n\Big]x^{n+2}=0$$

Como os monômios x^n são linearmente independentes, segue que

$$[\lambda_2-m(m+1)]a_0+2a_2=0 \tag{344}$$

$$[\lambda_2-(2+m(m+3))]a_1+6a_3=0 \tag{345}$$

e também, para todo $n\geq 0$,

$$a_{n+4}=\frac{[m(m+1)-\lambda_2+2(n+2)(n+m+2)]a_{n+2}-[m(m+1)-\lambda_2+n(n+2m+1)]a_n}{(n+4)(n+3)} \tag{346}$$

que é a relação de recorrência.

A equação (344) nos diz que

$$a_2=\frac{m(m+1)-\lambda_2}{2\cdot 1}a_0$$

e (345) implica

$$a_3=\frac{2+m(m+3)-\lambda_2}{3\cdot 2}a_1$$

Queremos encontrar uma relação de recorrência mais amigável que envolva apenas a_{n+2} e a_n. As duas últimas expressões nos sugerem tentar

$$a_{n+2}=\frac{c_n+m(m+d_n)-\lambda_2}{(n+2)(n+1)}a_n$$

onde c_n e d_n são coeficientes a determinar. Rigorosamente, esses termos podem depender de m ou de λ_2. Mas vamos descartar essa hipótese por hora.

Do apêndice (I), sabemos que para $m = 0$ a relação precisa ser

$$a_{n+2} = \frac{n(n+1) - \lambda_2}{(n+2)(n+1)} a_n$$

o que nos leva a fazer $c_n = n(n+1)$. Então para qualquer m deve ficar

$$a_{n+2} = \frac{n(n+1) + m(m + d_n) - \lambda_2}{(n+2)(n+1)} a_n$$

Isolando d_n, obtemos

$$d_n = \left((n+2)(n+1) \frac{a_{n+2}}{a_n} + \lambda_2 - n(n+1) \right) \cdot \frac{1}{m} - m \qquad (347)$$

Como já temos uma relação de recorrência de a_{n+4} em função de a_{n+2} e a_n, podemos encontrar todos os coeficientes d_n a partir dos valores conhecidos de a_n. Com isso, podemos depois tentar encontrar uma expressão para d_n se tiver um padrão fácil de perceber.

Se a hipótese de c_n e d_n dependerem apenas de n estiver correta, a equação (347) deve retornar os mesmos valores independentemente do valor de m e de λ_2. Usando $m = 1$, $\lambda_2 = 0$, $a_0 = 1$ e $a_1 = 0$, a relação (347) vai retornar os coeficientes pares de d_n:

$$d_0 = 1 \quad d_2 = 5 \quad d_4 = 9 \quad d_6 = 13 \quad d_8 = 17$$

E usando os mesmos parâmetros mas com $a_0 = 0$ e $a_1 = 1$, a expressão (347) retornará

$$d_1 = 3 \quad d_3 = 7 \quad d_5 = 11 \quad d_7 = 15 \quad d_9 = 19$$

Esse padrão nos sugere

$$d_n = 2n + 1$$

de forma que

$$a_{n+2} = \frac{n(n+1) + m(m + 2n + 1) - \lambda_2}{(n+2)(n+1)} a_n \qquad (348)$$

Para verificarmos se essa relação está correta, vamos isolar a_n para obter

$$a_n = \frac{(n+2)(n+1)}{n(n+1) + m(m + 2n + 1) - \lambda_2} a_{n+2}$$

Inserindo essa expressão na relação de recorrência (346) que já conhecemos e notando que

$$m(m+1) - \lambda_2 + n(n + 2m + 1) = m^2 + m - \lambda_2 + n^2 + 2mn + n \qquad (349)$$

$$= n(n+1) + m(m + 2n + 1) - \lambda_2 \qquad (350)$$

obteremos

$$\begin{aligned}(n+4)(n+3)a_{n+4} &= [m(m+1)-\lambda_2+2(n+2)(n+m+2)]a_{n+2}-(n+2)(n+1)a_{n+2}\\ &= [m(m+1)-\lambda_2+2(n+2)(n+m+2)-(n+2)(n+1)]\,a_{n+2}\\ &= [m(m+1)-\lambda_2+(n+2)\cdot[2(n+m+2)-(n+1)]]\,a_{n+2}\\ &= [m(m+1)-\lambda_2+(n+2)\cdot[n+2m+3]]\,a_{n+2}\end{aligned}$$

e então

$$a_{n+4} = \frac{m(m+1)-\lambda_2+(n+2)\cdot[n+2m+3]}{(n+4)(n+3)}a_{n+2}$$

Reduzindo o índice de $n+4$ para $n+2$, obteremos

$$a_{n+2} = \frac{m(m+1)-\lambda_2+n\cdot[n+2m+1]}{(n+2)(n+1)}a_n$$

Usando novamente a relação (350), chegamos na expressão

$$a_{n+2} = \frac{n(n+1)+m(m+2n+1)-\lambda_2}{(n+2)(n+1)}a_n \tag{351}$$

que é exatamente a relação (348) que supomos ser e, consequentemente, a relação foi provada por indução.

Quando n é muito grande e lembrando que m é não negativo, a relação se comportará como

$$a_{n+2} \approx \frac{n(n+1)+2nm}{(n+2)(n+1)}a_n > \frac{n(n+1)}{(n+2)(n+1)}a_n = \frac{n}{n+2}a_n$$

Do apêndice (I), sabemos que a relação

$$\frac{n}{n+2}a_n$$

faz a série de potência $\sum a_n x^n$ divergir quando $x = \pm 1$. Pelo teste da comparação (2.9), essa série com $m > 0$ também vai divergir. Lembre-se que supomos inicialmente que as séries das funções

$$f_{1_1}(x) = |1-x|^{m/2}\sum_{n=0}^{\infty} e_n(x-1)^n \tag{352}$$

e

$$f_{1_{-1}}(x) = |1+x|^{m/2}\sum_{n=0}^{\infty} g_n(x+1)^n \tag{353}$$

não divergiam em nenhum ponto do intervalo $[-1,1]$ e portanto podíamos rearranjar os termos $(1-x)^n$ e $(1+x)^n$ dentro da soma infinita para encontrar os coeficientes a_n relacionados aos monômios x^n. Se divergissem, não poderíamos fazer isso e a solução teria de ser encontrada substituindo (352) ou (353) na equação diferencial de Legendre associada.

O método de Frobenius garante que $f_{1_1}(x)$ converge em $(-1, 2)$ e que $f_{1_{-1}}(x)$ converge em $(-2, 1)$. Então nada sabemos sobre a convergência de $f_{1_{-1}}$ em $x = 1$ e de f_{1_1} em $x = -1$. Então precisamos truncar a série $\sum a_n x^n$ se quisermos validar a solução

$$f_1(x) = (1 - x^2)^{m/2} \sum_{n=0}^{\infty} a_n x^n$$

no intervalo $[-1, 1]$. Verificando a relação de recorrência (351), podemos observar que isso só pode ser possível se

$$n(n + 1) + m(m + 2n + 1) - \lambda_2 = 0$$

para algum $n = j$ não negativo. Isso significa que devemos escolher a constante de separação λ_2 de forma que

$$\lambda_2 = j(j + 1) + m(m + 2j + 1)$$

Distribuindo os produtos, obteremos

$$\begin{aligned} \lambda_2 &= j^2 + j + m^2 + 2mj + m \\ &= (j + m)^2 + (j + m) \end{aligned}$$

Denotando convenientemente $j + m = l$, temos que

$$\lambda_2 = l^2 + l = l(l + 1)$$

que é a mesma forma que encontramos no apêndice (I) para $m = 0$.

Agora uma observação importante: como $l = m + j$ para algum j não negativo, segue que é necessário restringir o inteiro l à condição

$$l \geq m$$

caso contrário toda a argumentação usada para truncar a série $\sum a_n x^n$ não será mais válida. De fato, podemos reescrever a relação de recorrência como

$$a_{n+2} = \frac{(n + m)(n + m + 1) - l(l + 1)}{(n + 2)(n + 1)} a_n \tag{354}$$

Se $l < m$, de forma que $l = m - u$ para algum u positivo, então

$$a_{n+2} = \frac{(n + m)(n + m + 1) - (m - u)(m - u + 1)}{(n + 2)(n + 1)} a_n$$

o que significa que a série só poderá ser truncada se $n + m = m - u$, ou $n = -u$, o que está fora de questão. Isso implica na divergência da solução em $x = \pm 1$.

Também note que, da relação de recorrência, os coeficientes pares a_{2n} só dependem de a_0 enquanto os ímpares de a_1, de forma que podemos escrever

$$\sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} a_{2n} x^{2n} + \sum_{n=0}^{\infty} a_{2n+1} x^{2n+1}$$

Além disso, o truncamento só ocorrerá para um dos somatórios, o que implica que devemos fazer $a_0 = 0$ ou $a_1 = 0$ conforme o valor de $\lambda_2 = l(l+1)$. Revendo a relação de recorrência (354), o truncamento acontecerá quando $n = j$. Como $m+j = l$, temos $j = l-m$, o que significa que o polinômio que resultará da série $\sum a_n x^n$ terá grau $l-m$. Se $l-m$ for um número par, devemos fazer $a_1 = 0$. Analogamente, devemos impor $a_0 = 0$ se $l-m$ for ímpar.

Finalmente, as soluções da EDO

$$(1-x^2)\Theta''(x) - 2x\Theta'(x) + \left[l(l+1) - \frac{m^2}{1-x^2}\right]\Theta(x) = 0 \tag{355}$$

que não divergem em $[-1,1]$ são denotadas por $\Theta(x) = P_l^m(x)$ e dadas por

$$P_l^m(x) = (1-x^2)^{m/2}\left(\sum_{n=0}^{\infty} a_{2n}x^{2n} + \sum_{n=0}^{\infty} a_{2n+1}x^{2n+1}\right)$$

com ou a_0 ou a_1 nulo conforme o caso. São denominadas de funções de Legendre associadas do primeiro tipo.

As outras soluções linearmente independentes da EDO (355) serão as dadas pelo teorema de Frobenius (6.1). Essas soluções são denominadas de funções de Legendre associadas do segundo tipo e são denotadas por $\Theta(x) = Q_l^m(x)$.

A equação diferencial (355) nos diz que a solução deve ser invariante sob o sinal de m, de forma que se $m < 0$, devemos trocar m por $-m$ em todas as equações, resultando em uma solução linearmente dependente em relação a solução com m positivo.

De maneira análoga, a EDO (355) é invariante se trocarmos l por $-(l+1)$, resultando no mesmo valor $\lambda_2 = l(l+1)$, de forma que

$$P_l^m = P_{-(l+1)}^m$$

Se definirmos $l' = l+1$, teremos

$$P_{l'-1}^m = P_{-l'}^m \implies P_{-l}^m = P_{l-1}^m$$

As cinco primeiras funções associadas de Legendre normalizadas (para $m = 1$, $m = 2$ e $m = 5$) são dadas nas figuras (10),(11) e (12). Para isso, foi necessário três programas em Python: um para gerar os coeficientes da série de Taylor dessas funções usando a relação de recorrência (351) (no apêndice (I.2.1)); um para gerar as normas

$$\|P_l^m(x)\| = \sqrt{\int_{-1}^{1} |P_l^m(x)|^2 \, \mathrm{d}x}$$

das funções usando a regra de Simpson da integral (I.2.2); e um programa para imprimir várias dessas funções simultaneamente na tela (I.2.3).

Revendo a equação diferencial de Legendre associada

$$\frac{d}{dx}\left((1-x^2)\frac{d}{dx}P_l^m(x)\right) + \left[l(l+1) - \frac{m^2}{1-x^2}\right]P_l^m(x) = 0$$
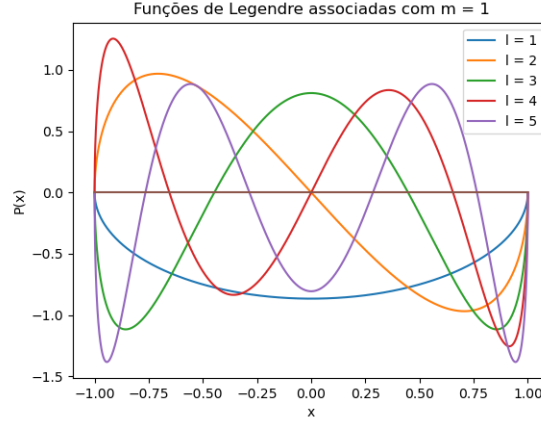
Figura 10: Primeiras cinco funções associadas de Legendre normalizadas com $m = 1$.
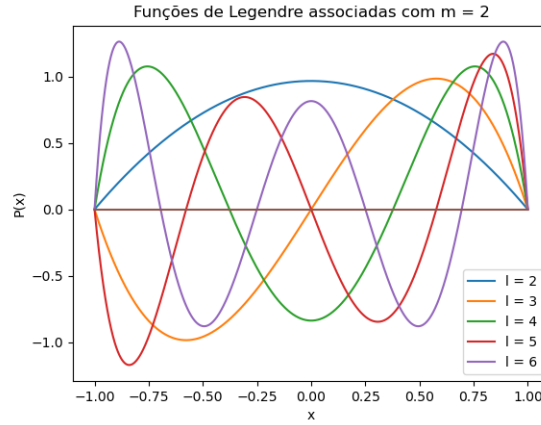
Figura 11: Primeiras cinco funções associadas de Legendre normalizadas com $m = 2$.

podemos fazer $p(x) = 1 - x^2$,

$$q_1(x) = l_1(l_1 + 1) - \frac{m^2}{1 - x^2} \quad \text{e} \quad q_1(x) = l_2(l_2 + 1) - \frac{m^2}{1 - x^2}$$

Agora suponha que $l_2 > l_1$. Isso implica $q_2 > q_1$ em todo o intervalo $I = (-1, 1)$. Como p é positivo em I, podemos usar o teorema da comparação de Sturm (4.3) para mostrar que entre duas raízes consecutivas de $P_l^m(x)$, haverá pelo menos uma raiz de $P_{l+1}^m(x), P_{l+2}^m(x), P_{l+3}^m(x), \ldots$ e todos os demais. Apesar de q_1 e
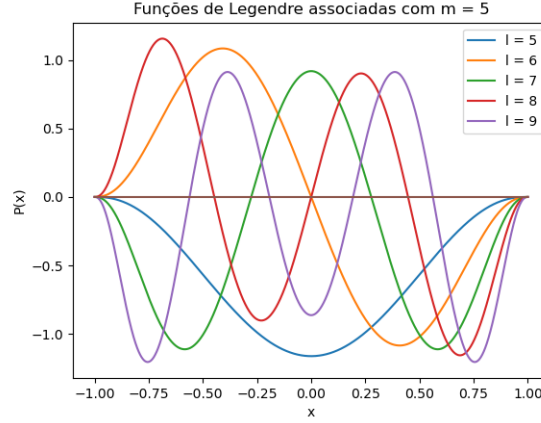
Figura 12: Primeiras cinco funções associadas de Legendre normalizadas com $m = 5$.

q_2 serem descontínuas em $x = \pm 1$, aparentemente essa propriedade se mantém para as raízes em $x = \pm 1$. Isso pode ser verificado nas três figuras: (10), (11) e (12).

Dois gráficos das funções de Legendre associadas com m variando e mantendo l fixo pode ser visto nas figuras (13) e (14).

Note que agora são as raízes de P_l^0, P_l^1, P_l^2, P_l^3, $\ldots$, P_l^{m-1} que se interpõem entre duas raízes consecutivas de P_l^m, comportamento invertido de quando l varia. Isso acontece porque os termos $l(l+1)$ e $\dfrac{m^2}{1-x^2}$ possuem sinais trocados na equação diferencial e consequentemente o teorema da comparação de Sturm se aplica às ordens inferiores a m.

6.3.1 Propriedades das funções de Legendre associadas

Agora note que se m for par, o fator $(1-x^2)^{m/2}$ será um polinômio de grau m, e então $P_l^m(x)$ será um polinômio de grau $(l-m)+m = l$.

E quando $l = m$, a série $\sum a_n x^n$ só terá o primeiro termo a_0, resultando que

$$P_l^l(x) = a_0(1-x^2)^{l/2}$$

o que significa que $P_l^l(x)$ terá apenas duas raízes, em $x = 1$ e $x = -1$. Se l for par, $P_l^l(x)$ será um polinômio e ambas as raízes $x = \pm 1$ terão multiplicidade $l/2$.

Como $(1-x^2)^{m/2}$ é sempre uma função par, segue que P_l^m será par ou ímpar a depender da soma $\sum a_n x^n$, que é ímpar se $l - m$ for ímpar e vice-versa. Como $x = \pm 1$ são as duas raízes consecutivas de $P_l^l(x)$, segue que as funções P_l^0, P_l^1, P_l^2, $\ldots$, P_l^{l-1} terão pelo menos uma raiz no intervalo aberto $(-1, 1)$, o que podemos ver na figura (13).
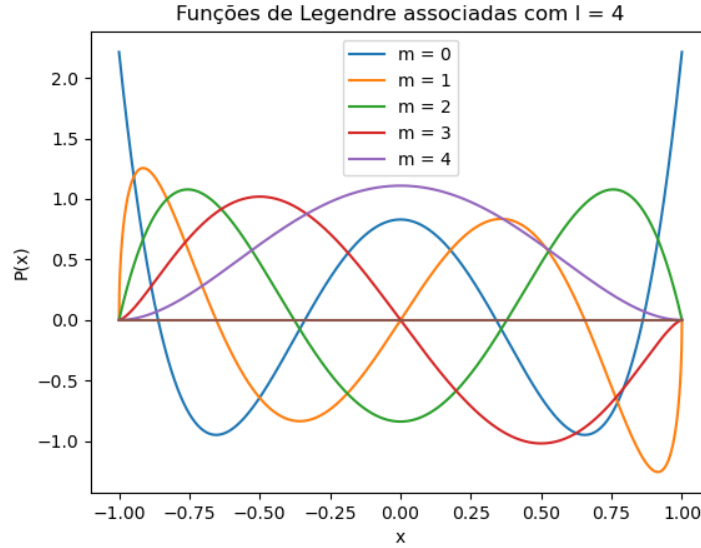
Figura 13: Primeiras funções de Legendre associadas com $l = 4$ e m variando de 0 a 4.

Podemos explorar mais ainda o teorema da comparação de Sturm: como P_l^{l-1} é ímpar e contínua em $[-1, 1]$, segue que P_l^{l-1} terá uma raiz em $x = 0$. Isso implica que as funções $P_l^0,\ P_l^1,\ P_l^2,\ \dots,\ P_l^{l-2}$ terão pelo menos uma raiz em $(-1, 0)$ e em $(0, 1)$. Como $x = \pm 1$ sempre será raiz de todas as funções P_l^m com $m \geq 1$ e P_l^{l-3} é ímpar e contínua, então as funções P_l^{l-4} e com ordens inferiores terão pelo menos duas raízes em $(-1, 0)$ e em $(0, 1)$.

Agora suponha um intervalo I cujas extremidades são duas raízes consecutivas de P_l^m, com $m \geq 1$. Quantas raízes a função P_l^{m-1} terá em I? O teorema de Sturm garante que existe pelo menos uma, mas a priori podem existir mais. Sabemos que a série $\sum a_n x^n$ será um polinômio de grau $l - m$, o que significa que em toda a reta real esse polinômio terá no máximo $l - m$ raízes. Sabemos que P_l^l não tem raízes em $(-1, 1)$, mas P_l^{l-1} precisa ter pelo menos uma (que é $x = 0$). Podem existir outras em $(-1, 1)$? Não porque o polinômio $\sum a_n x^n$ terá grau 1 e consequentemente só pode ter uma raiz. As outras duas de P_l^{l-1} estão nas extremidades. Analogamente, P_l^{l-2} só pode ter duas raízes em $(-1, 1)$ porque $\sum a_n x^n$ terá grau 2 e ambas precisam satisfazer o teorema de Sturm, sendo cada uma no intervalo $(-1, 0)$ e em $(0, 1)$. Podemos iterar essa argumentação para mostrar que $\sum a_n x^n$ terá $l - m$ raízes em $(-1, 1)$ e consequentemente P_l^m terá $2 + l - m$ raízes em $[-1, 1]$ (com as outras duas sendo as extremidades). Tudo isso também pode ser visto nas figuras (13) e (14). Conforme veremos adiante, isso terá implicações nos harmônicos esféricos e na função de onda ψ.

Uma outra propriedade das funções de Legendre associadas é que se $m = 1$,
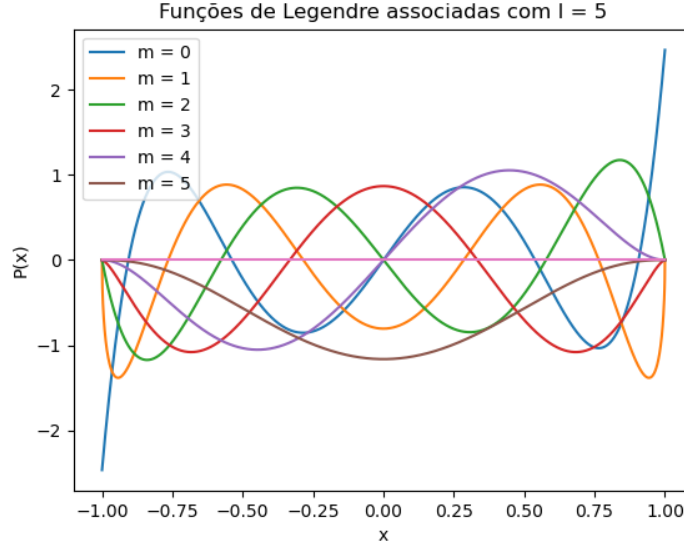
Figura 14: Primeiras funções de Legendre associadas com $l = 5$ e m variando de 0 a 5.

então a primeira derivada de $P_l^1(x)$ terá um fator $(1-x^2)^{-1/2}$, no que implica em uma divergência em $x = \pm 1$. Podemos ver nitidamente as duas tangentes verticais na figura (10) em $x = 1$ e $x = -1$. Isso não acontece para as demais ordens m porque as derivadas terão uma potência positiva de $(1-x^2)$.

O próximo teorema nos permite relacionar P_l^m com P_l^0 usando os coeficientes a_n que encontramos anteriormente.

Teorema 6.2. *A série* $\displaystyle\sum_{n=0}^{\infty} a_n x^n$ *é um múltiplo de* $\dfrac{d^m}{dx^m}P_l(x)$*, onde* $P_l(x) = P_l^0(x)$.

Demonstração. Seja $f(x) = \displaystyle\sum_{n=0}^{\infty} b_n x^n$. A m-ésima derivada de $f(x)$ é dada por

$$
\begin{aligned}
\frac{d^m}{dx^m} f(x) &= \sum_{n=0}^{\infty} n(n-1)(n-2)\cdot \ldots \cdot (n-m+1) b_n x^{n-m} \\
&= \sum_{n=m}^{\infty} \frac{n!}{(n-m)!} b_n x^{n-m} \\
&= \sum_{n=0}^{\infty} \frac{(n+m)!}{n!} b_{n+m} x^n
\end{aligned}
$$

Agora vamos fazer $a_n = \dfrac{(n+m)!}{n!} b_{n+m}$ e supor que b_n satisfaz a relação de recorrência

$$b_{n+2} = R_n^b b_n \tag{356}$$

enquanto a_n satisfaz

$$a_{n+2} = R_n^a a_n \tag{357}$$

Então

$$\begin{aligned} a_{n+2} &= R_n^a a_n \\ \frac{(n+m+2)!}{(n+2)!} b_{n+m+2} &= R_n^a \frac{(n+m)!}{n!} b_{n+m} \\ \frac{(n+m+2)!}{(n+2)!} R_{n+m}^b b_{n+m} &= R_n^a \frac{(n+m)!}{n!} b_{n+m} \end{aligned}$$

no que implica

$$\begin{aligned} R_{n+m}^b &= \frac{(n+m)!}{n!} \cdot \frac{(n+2)!}{(n+m+2)!} R_n^a \\ &= \frac{(n+1)(n+2)}{(n+m+1)(n+m+2)} R_n^a \end{aligned}$$

Substituindo R_n^a pela relação de recorrência (354), obteremos

$$\begin{aligned} R_{n+m}^b &= \frac{(n+1)(n+2)}{(n+m+1)(n+m+2)} \cdot \frac{(n+m)(n+m+1) - l(l+1)}{(n+2)(n+1)} \\ &= \frac{(n+m)(n+m+1) - l(l+1)}{(n+m+1)(n+m+2)} \end{aligned}$$

Reduzindo o índice de $n+m$ para n, encontraremos a relação de recorrência de b_n,

$$R_n^b = \frac{n(n+1) - l(l+1)}{(n+1)(n+2)}$$

que é exatamente a relação (354) com $m = 0$ e portanto $f(x) = P_l(x)$. ∎

Isso significa que podemos reescrever $P_l^m(x)$ como sendo

$$P_l^m(x) = c_0 \cdot (1-x^2)^{m/2} \frac{d^m}{dx^m} P_l(x) \tag{358}$$

com uma constante c_0 arbitrária.

Usando o seguinte teorema, poderemos encontrar uma outra forma equivalente para (358).

Teorema 6.3. *A função $P_l(x) = P_l^0(x)$ é um múltiplo de $\dfrac{d^l}{dx^l}(1-x^2)^l$.*

Demonstração. Seja $f(x) = \sum_{n=0}^{\infty} b_n x^n$. Podemos repetir os passos da demonstração do teorema (6.2), porém agora inserindo

$$R_n^a = \frac{n(n+1) - l(l+1)}{(n+2)(n+1)}$$

e trocando a letra m por l na equação

$$R_{n+l}^b = \frac{(n+1)(n+2)}{(n+l+1)(n+l+2)} R_n^a$$

no que resulta

$$\begin{aligned} R_{n+l}^b &= \frac{(n+1)(n+2)}{(n+l+1)(n+l+2)} \cdot \frac{n(n+1) - l(l+1)}{(n+2)(n+1)} \\ &= \frac{n(n+1) - l(l+1)}{(n+l+1)(n+l+2)} \end{aligned}$$

Reduzindo o índice de $n+l$ para n, obteremos

$$\begin{aligned} R_n^b &= \frac{(n-l)(n-l+1) - l(l+1)}{(n+1)(n+2)} \\ &= \frac{n^2 + n - nl - nl + l^2 - l}{(n+1)(n+2)} \\ &= \frac{n(n+1) - 2nl - 2l}{(n+1)(n+2)} \\ &= \frac{n(n+1) - 2l(n+1)}{(n+1)(n+2)} \\ &= \frac{(n+1)(n-2l)}{(n+1)(n+2)} \\ &= \frac{n-2l}{n+2} \end{aligned}$$

e então

$$b_{n+2} = \frac{n-2l}{n+2} b_n$$

Essa relação nos diz que a série $\sum b_n x^n$ truncará no coeficiente a_{2l}, ou seja, para um n par. Para os coeficientes pares, temos

$$b_{2n+2} = \frac{2n-2l}{2n+2} b_{2n} = \frac{n-l}{n+1} b_{2n} \tag{359}$$

Os primeiros coeficientes pares são

$$b_2 = -lb_0 \quad b_4 = \frac{1-l}{2} b_2 = \frac{-(l-1)}{2} \cdot -lb_0 = \frac{(-1)^2 l(l-1)}{2} b_0$$

$$b_6 = \frac{-(l-2)}{3} b_4 = \frac{(-1)^3(l-2)(l-1)l}{3 \cdot 2} b_0 \quad b_8 = \frac{(-1)^4(l-3)(l-2)(l-1)l}{4 \cdot 3 \cdot 2} b_0$$

Esse padrão nos sugere

$$b_{2n} = (-1)^n \frac{l!}{n!(l-n)!} b_0 \tag{360}$$

Inserindo essa expressão na relação de recorrência (359), obteremos

$$\begin{aligned} b_{2n+2} &= \frac{n-l}{n+1}(-1)^n \frac{l!}{n!(l-n)!} b_0 \\ &= (-1)^{n+1}(l-n)\frac{l!}{n!(n+1)(l-n)!} b_0 \\ &= (-1)^{n+1} \frac{l!}{(n+1)!(l-(n+1))!} b_0 \end{aligned}$$

que é (360) trocando n por $n+1$, o que demonstra a expressão.

Para os coeficientes ímpares, porém, a relação de recorrência é

$$b_{2n+3} = \frac{2(n-l)+1}{2n+3} b_{2n+1}$$

na qual não trunca para nenhum n e possui um comportamento assintótico $b_{2n+3} \approx b_{2n+1}$ para n grande. Sabemos que a l-ésima derivada de $f(x)$ produz $P_l(x)$, que é um polinômio e, consequentemente, tem finitos coeficientes não-nulos. Então iremos descartar os coeficientes ímpares fazendo $b_1 = 0$.

Como a série $\sum b_n x^n$ trunca no coeficiente a_{2l}, temos

$$f(x) = \sum_{n=0}^{l} b_{2n} x^{2n} = b_0 \sum_{n=0}^{l} (-1)^n \frac{l!}{n!(l-n)!} x^{2n}$$

A fração que está dentro do somatório é o coeficiente binomial, denotado por

$$\binom{l}{n} = \frac{l!}{n!(l-n)!}$$

e então

$$f(x) = b_0 \sum_{n=0}^{l} (-1)^n \binom{l}{n} x^{2n} = b_0 \sum_{n=0}^{l} (-1)^n \binom{l}{n} (x^2)^n = b_0 \sum_{n=0}^{l} \binom{l}{n} (-x^2)^n$$

Agora compare essa expressão com o teorema binomial,

$$(1+y)^l = \sum_{n=0}^{l} \binom{l}{n} y^n$$

trocando y por $-x^2$. Isso significa que

$$f(x) = b_0(1-x^2)^l$$

ou seja,

$$P_l(x) = \frac{d^l}{dx^l} f(x) = b_0 \frac{d^l}{dx^l}(1-x^2)^l$$

■

Na literatura, o teorema (6.3) aparece como

$$P_l(x) = \frac{1}{2^l l!} \frac{d^l}{dx^l}(1-x^2)^l \tag{361}$$

A expressão (361) é a fórmula de Rodrigues para os polinômios de Legendre.

Isso também implica que podemos reescrever as funções associadas de Legendre $P_l^m(x)$ como função das derivadas de $(1-x^2)^l$. Ou seja,

$$P_l^m(x) = (-1)^m \frac{1}{2^l l!}(1-x^2)^{m/2} \frac{d^{l+m}}{dx^{l+m}}(1-x^2)^l \tag{362}$$

onde o fator $(-1)^m$ é a fase de Condon–Shortley, que auxilia algumas operações em mecânica quântica, como os operadores escada.

Teorema 6.4 (Ortogonalidade e normalização). *Sejam* $P_l^m(x)$ *as funções dadas pela expressão* (362)*. Então*

$$\int_{-1}^{1} P_l^m(x) P_{l'}^m(x)\, \mathrm{d}x = \frac{2}{(2l+1)} \cdot \frac{(l+m)!}{(l-m)!} \delta_{ll'}$$

Demonstração. As funções $P_l^m(x)$ satisfazem o problema de Sturm-Liouville

$$\frac{d}{dx}\left(p(x)\frac{d}{dx}P_l^m(x)\right) + \left[q(x) + r(x)\lambda\right] P_l^m(x) = 0$$

com $p(x) = 1-x^2$, $q(x) = -\dfrac{m^2}{1-x^2}$, $r(x) = 1$ e $\lambda = l(l+1)$, com as condições de contorno

$$P_l^m(\pm 1) = (\pm 1)^l \delta_{0m} \tag{363}$$

O operador diferencial $L = \dfrac{d}{dx}\left(p(x)\dfrac{d}{dx}\right) + q(x)$ será hermitiano se

$$p[\overline{f'}g - \overline{f}g']\Big|_{-1}^{1} = 0$$

para quaisquer funções f e g contínuas que satisfaçam as condições de contorno (363). Como $p(x)$ se anula em $x = \pm 1$, isso garante que L seja hermitiano, independentemente se $m = 0$ ou $m \neq 0$.

Portanto, pelo teorema (4.12), as funções $P_l^m(x)$ satisfazem uma relação de ortogonalidade em relação ao autovalor $\lambda = l(l+1)$,

$$\int_{-1}^{1} P_l^m(x) P_{l'}^m(x)\, \mathrm{d}x = A_{lm}\delta_{ll'}$$

para alguma coeficiente A_{lm}.

O restante da demonstração, que mostra que

$$A_{lm} = \frac{2}{(2l+1)} \cdot \frac{(l+m)!}{(l-m)!}$$

está no apêndice (I.1) por ser demasiadamente longa. ■

A teoria de Sturm-Liouville também garante que o conjunto $\{P_0^m, P_1^m, P_2^m, \ldots\}$ forma uma base infinita do espaço $\mathcal{L}^2(-1,1)$ das funções quadrado integráveis no intervalo $[-1,1]$, com produto interno

$$\langle f, g \rangle = \int_{-1}^{1} f(x)g(x)r(x)\,\mathrm{d}x = \int_{-1}^{1} f(x)g(x)\,\mathrm{d}x$$

uma vez que $r(x) = 1$. E então

$$f(x) = \sum_{l=0}^{\infty} \langle \overline{P}_l^m, f \rangle\, \overline{P}_l^m(x) \quad \forall f \in \mathcal{L}^2(-1,1)$$

onde $\overline{P}_l^m$ são as funções associadas de Legendre normalizadas,

$$\overline{P}_l^m(x) = \frac{P_l^m(x)}{\|P_l^m\|} = \sqrt{\frac{1}{\langle P_l^m(x), P_l^m(x)\rangle}}\, P_l^m(x)$$

$$\overline{P}_l^m(x) = \sqrt{\frac{2l+1}{2} \cdot \frac{(l-m)!}{(l+m)!}}\, P_l^m(x)$$

Retornando à equação diferencial para o ângulo polar $\theta = \arccos(x)$, a solução é dada por

$$\Theta(\theta) = \overline{P}_l^m(\cos(\theta)) = \sqrt{\frac{2l+1}{2} \cdot \frac{(l-m)!}{(l+m)!}}\, P_l^m(\cos(\theta))$$

Usaremos a versão normalizada das funções associadas de Legendre porque será necessário para normalizar a função de onda $\psi(r,\theta,\varphi) = R(r)\Theta(\theta)\Phi(\varphi)$.

6.3.2 Soluções com 'm' não inteiro

Sabemos que a imposição m inteiro advém da condição de contorno periódica da função azimutal $\Phi(\varphi)$. Porém, a equação diferencial

$$\frac{d}{dx}\left((1-x^2)\Theta'(x)\right) + \left[\lambda_2 - \frac{m^2}{1-x^2}\right]\Theta(x) = 0 \tag{364}$$

tem soluções contínuas em $[-1,1]$ com m não inteiro. Agora a diferença das soluções α_1 e α_2 da equação indicial é um número não inteiro, mas isso não

afeta a forma da primeira solução linearmente independente (que é o caso não-inteiro do teorema de Frobenius), o que implica que a relação de recorrência

$$a_{n+2} = \frac{(n+m)(n+m+1) - \lambda_2}{(n+2)(n+1)} a_n$$

continua válida. Verificando-a, vemos que a série $\sum a_n x^n$ pode ser truncada se λ_2 tiver a forma

$$\lambda_2 = (j+m)(j+m+1)$$

e $a_0 = 0$ se j for ímpar e vice-versa. Nesse caso, o truncamento aconteceria quando $n = j$ e a relação de recorrência ficaria

$$a_{n+2} = \frac{(n+m)(n+m+1) - (j+m)(j+m+1)}{(n+2)(n+1)} a_n$$

Isso significa que não existe qualquer restrição para m se quisermos uma solução da EDO (364) que não divirja em $x = \pm 1$.

A equação diferencial ficaria, portanto,

$$\frac{d}{dx}\left((1-x^2)\frac{d}{dx}P_j^m(x)\right) + \left[(j+m)(j+m+1) - \frac{m^2}{1-x^2}\right]P_j^m = 0$$

Um exemplo é mostrado na figura (15), onde usamos $m = 1/5$. Note que quando $m \to 0^+$, as soluções começam a ter aparência de polinômios, com exceção nas bordas, onde o fator $(1-x^2)^{m/2}$ força-os a se anularem nesses pontos. Conforme podemos ver na figura (15), a primeira solução tende a ser uma reta horizontal (que é a função constante) enquanto que a segunda e a terceira solução tendem a ser uma reta e uma parábola respectivamente. Em outras palavras, $\forall x \in (-1, 1)$ as soluções começam a convergir pontualmente para os polinômios $P_l(x) = P_l^0(x)$ de Legendre, com $l = j$.

6.4 Equação de Bessel esférica

Para a equação diferencial da variável r (332), temos

$$r^2 R''(r) + 2rR'(r) + \left(\frac{2m_0 E}{\hbar^2} r^2 - \lambda_2\right) R(r) = 0 \tag{365}$$

Sabendo que $\lambda_2 = l(l+1)$ e definindo

$$k = \frac{\sqrt{2m_0 E}}{\hbar}$$

a equação (365) se torna

$$r^2 R''(r) + 2rR'(r) + \left(k^2 r^2 - l(l+1)\right) R(r) = 0 \tag{366}$$

Note que a energia E precisa ser não-negativa, já que o potencial mínimo é zero e o teorema (3.4) garante que a solução não será normalizável se $E < V_{min} = 0$.
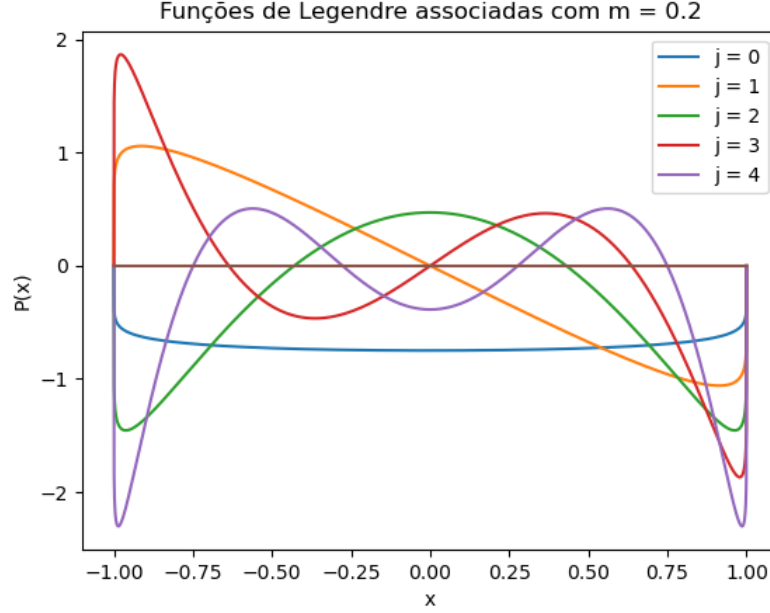
Figura 15: Soluções normalizadas da equação de Legendre associada com $m = 1/5$.

As duas soluções reais e linearmente independentes da equação diferencial (366) são denotadas por $j_l(x)$ e $y_l(x)$, denominadas funções de Bessel esféricas.

Dividindo (366) por r^2, temos a equação diferencial

$$R''(r) + \frac{2}{r}R'(r) + \left(k^2 - \frac{l(l+1)}{r^2}\right)R(r) = 0$$

As funções $\frac{2}{r}$ e $k^2 + \frac{l(l+1)}{r^2}$ são analíticas em intervalos que não incluam $r = 0$. Entretanto, se usarmos o método padrão de série de potências, não teremos certeza se existem duas soluções analíticas centradas em $r = 0$.

Devemos notar que

$$p_0 = \lim_{r\to 0} r \cdot \frac{2}{r} = \lim_{r\to 0} 2 = 2 \in \mathbb{R}$$

$$q_0 = \lim_{r\to 0} r^2 \cdot \left(k^2 - \frac{l(l+1)}{r^2}\right) = \lim_{r\to 0} r^2k^2 - l(l+1) = -l(l+1) \in \mathbb{R}$$

Como os limites existem e são finitos, então $x = 0$ é um ponto singular regular e podemos usar o método de Frobenius.

Substituindo p_0 e q_0 na equação indicial, obteremos a equação

$$\begin{aligned} 0 &= \alpha(\alpha-1) + 2\alpha - l(l+1) \\ &= \alpha(\alpha+1) - l(l+1) \end{aligned}$$

Suas soluções são

$$\alpha_1 = l \quad \alpha_2 = -(l+1)$$

Nesse caso, $\alpha_1 - \alpha_2 = l + (l+1) = 2l+1$, que é um inteiro não-negativo e então podemos aplicar o teorema (6.1).

Para a primeira solução $f_1(r)$, fazemos

$$f_1(r) = r^\alpha \sum_{n=0}^{\infty} a_n r^n = \sum_{n=0}^{\infty} a_n r^{n+\alpha}$$

$$f_1'(r) = \sum_{n=0}^{\infty} (n+\alpha) a_n r^{n+\alpha-1} = r^{\alpha-1} \sum_{n=0}^{\infty} (n+\alpha) a_n r^n$$

$$f_1''(r) = \sum_{n=0}^{\infty} (n+\alpha)(n+\alpha-1) a_n r^{n+\alpha-2} = r^{\alpha-2} \sum_{n=0}^{\infty} (n+\alpha)(n+\alpha-1) a_n r^n$$

Substituindo em (366), temos

$$r^\alpha \sum_{n=0}^{\infty} (n+\alpha)(n+\alpha-1) a_n r^n + r^\alpha \sum_{n=0}^{\infty} 2(n+\alpha) a_n r^n + \\ r^\alpha \sum_{n=0}^{\infty} k^2 a_n r^{n+2} - r^\alpha \sum_{n=0}^{\infty} l(l+1) a_n r^n = 0$$

$$r^\alpha \left[\sum_{n=0}^{\infty} \left[(n+\alpha)(n+\alpha-1) a_n + 2(n+\alpha) a_n - l(l+1) a_n \right] r^n + \sum_{n=0}^{\infty} k^2 a_n r^{n+2} \right] = 0$$

A última equação é válida mesmo para $r \neq 0$ e portanto

$$\sum_{n=0}^{\infty} \left[(n+\alpha)(n+\alpha-1) a_n + 2(n+\alpha) a_n - l(l+1) a_n \right] r^n + \sum_{n=0}^{\infty} k^2 a_n r^{n+2} = 0 \tag{367}$$

Do primeiro somatório:

$$\begin{aligned} \left[(n+\alpha)(n+\alpha-1) + 2(n+\alpha) - l(l+1) \right] a_n &= \left[(n+\alpha) \cdot \left[(n+\alpha-1) + 2 \right] - l(l+1) \right] a_n \\ &= \left[(n+\alpha)(n+\alpha+1) - l(l+1) \right] a_n \end{aligned}$$

e também

$$\sum_{n=0}^{\infty} \left[(n+\alpha)(n+\alpha+1) - l(l+1) \right] a_n r^n =$$

$$[\alpha(\alpha+1) - l(l+1)] a_0 + \left[(\alpha+1)(\alpha+2) - l(l+1) \right] a_1 r + \sum_{n=2}^{\infty} \left[(n+\alpha)(n+\alpha+1) - l(l+1) \right] a_n r^n$$

com

$$\sum_{n=2}^{\infty}\left[(n+\alpha)(n+\alpha+1)-l(l+1)\right]a_n r^n=\sum_{n=0}^{\infty}\left[(n+\alpha+2)(n+\alpha+3)-l(l+1)\right]a_{n+2}r^{n+2}$$

Substituindo em (367), obtemos

$$\begin{gathered}[\alpha(\alpha+1)-l(l+1)]a_0r^0+\left[(\alpha+1)(\alpha+2)-l(l+1)\right]a_1r^1+\\ \sum_{n=0}^{\infty}\Bigg[\left[(n+\alpha+2)(n+\alpha+3)-l(l+1)\right]a_{n+2}+k^2a_n\Bigg]r^{n+2}=0\end{gathered}$$

Como os monômios r^n são linearmente independentes e a equação é válida para qualquer r, segue que

$$[\alpha(\alpha+1)-l(l+1)]a_0=0 \tag{368}$$

$$\left[(\alpha+1)(\alpha+2)-l(l+1)\right]a_1=0 \tag{369}$$

e para todo $n\geq 0$

$$\left[(n+\alpha+2)(n+\alpha+3)-l(l+1)\right]a_{n+2}+k^2a_n=0 \tag{370}$$

A equação (368) é a indicial. Nessa solução, $\alpha=\alpha_1=l$, o que significa que a_0 pode assumir qualquer valor enquanto que l não é solução de (369), o que implica na escolha $a_1=0$. Já a_0 é arbitrário.

A relação de recorrência é, portanto,

$$a_{n+2}=-\frac{k^2a_n}{(n+l+2)(n+l+3)-l(l+1)}\quad\forall n\geq 0 \tag{371}$$

Agora note que

$$\begin{aligned}(n+l+2)(n+l+3)-l(l+1)&=n^2+2nl+l^2+5n+5l+6-l^2-l\\&=n^2+2nl+5n+4l+6\\&=n^2+n\cdot(2l+5)+(2l+5)+2l+1\\&=n^2+(2l+5)\cdot(n+1)+2l+1\\&=n^2+(2l+5)\cdot(n+1)+(2l+5)-4\\&=(n^2-4)+(2l+5)\cdot[(n+1)+1]\\&=(n-2)\cdot(n+2)+(2l+5)\cdot(n+2)\\&=(n+2)\cdot[(2l+5)+(n-2)]\\&=(n+2)\cdot(n+3+2l)\end{aligned}$$

e então

$$a_{n+2}=-\frac{k^2a_n}{(n+2)\cdot(n+3+2l)}\quad\forall n\geq 0 \tag{372}$$

Como $a_1 = 0$, todos os coeficientes ímpares serão nulos. E para os pares,

$$a_{2n+2} = -\frac{k^2 a_{2n}}{(2n+2)\cdot(2n+3+2l)} \quad \forall n \geq 0 \tag{373}$$

Os primeiros termos não-nulos são

$$a_2 = \frac{-k^2 a_0}{6+4l} = \frac{-k^2 a_0}{2(3+2l)} \quad a_4 = \frac{-k^2 a_2}{20+8l} = \frac{k^4 a_0}{2^2 2(5+2l)\cdot(3+2l)}$$

$$a_6 = \frac{-k^6 a_0}{2^3 2\cdot 3(7+2l)\cdot(5+2l)\cdot(3+2l)}$$

$$a_8 = \frac{k^8 a_0}{2^4 2\cdot 3\cdot 4(9+2l)\cdot(7+2l)\cdot(5+2l)\cdot(3+2l)}$$

Esses termos sugerem que

$$a_{2n} = \frac{a_0(-1)^n k^{2n}}{2^n n!}\cdot\prod_{\alpha=0}^{n-1}\frac{1}{2\alpha+3+2l} \quad \forall n \geq 1 \tag{374}$$

Ao substituir (374) em (373), temos

$$\begin{aligned} a_{2(n+1)} = a_{2n+2} &= \frac{(-1)k^2}{(2n+2)(2n+3+2l)}\cdot\frac{a_0(-1)^n k^{2n}}{2^n n!}\cdot\prod_{\alpha=0}^{n-1}\frac{1}{2\alpha+3+2l} \\ &= \frac{a_0(-1)^{n+1}k^{2n+2}}{2(n+1)2^n n!}\cdot\prod_{\alpha=0}^{n}\frac{1}{2\alpha+3+2l} \\ &= \frac{a_0(-1)^{n+1}k^{2n+2}}{2^{n+1}(n+1)!}\cdot\prod_{\alpha=0}^{n}\frac{1}{2\alpha+3+2l} \end{aligned}$$

que é (374) substituindo $2n$ por $2(n+1)$ e portanto provamos a expressão (374) por indução.

O produtório em (374) também pode ser escrito na forma

$$\prod_{\alpha=0}^{n-1}\frac{1}{3+2l+2\alpha} = \frac{(1+2l)!!}{(3+2l+2(n-1))!!} = \frac{(1+2l)!!}{(1+2l+2n))!!}$$

e então

$$a_{2n} = \frac{a_0(-1)^n k^{2n}}{2^n n!}\cdot\frac{(1+2l)!!}{(1+2l+2n))!!} \quad \forall n \geq 1$$

Note que a constante k^{2n} pode ser absorvida pela potência r^{2n}, de modo que podemos definir uma outra função $R_2(kr) = R(r)$ que satisfaz a mesma equação diferencial (366) mas com $k = 1$. Essa função é a de Bessel esférica j_l, com $j_l(kr) = R_2(kr)$.

As primeiras soluções $f_1(r) = j_l(kr)$ com $k = 1$ são mostradas na figura (16). O resultado foi obtido usando três programas em Python: um para encontrar os

coeficientes usando a relação de recorrência (373)(código no apêndice (H.1.1)); outro para encontrar as normas (375) em relação a um intervalo infinito (H.1.2); e um terceiro para imprimir várias funções esféricas de Bessel na tela simultaneamente (H.1.3).

$$\|j_l(r)\| = \sqrt{\int_0^{\infty} [j_l(r)]^2 r^2 \, \mathrm{d}r} \tag{375}$$

Agora note que as raízes de j_l estão entre duas raízes de j_{l+1}. Revendo a equação diferencial

$$r^2 R''(r) + 2rR'(r) + [k^2 r^2 - l(l+1)]R(r) = 0$$

podemos fazer $p(r) = r^2$, $q_1(r) = k^2 r^2 - l_1(l_1+1)$ e $q_2(r) = k^2 r^2 - l_2(l_2+1)$. Se $l_2 > l_1$, então $q_1(r) > q_2(r)$ em todo o intervalo $[0, +\infty)$, enquanto que $p(r)$ é não negativo e só se anula em $r = 0$. Essas condições satisfazem os requisitos do teorema da comparação de Sturm (4.3), o que implica haverá pelo menos uma raiz de j_{l-1}, j_{l-2}, $j_{l-3}, \ldots,$ j_0 entre duas raízes consecutivas de j_l. Isso pode ser visto com mais detalhe na figura (17).
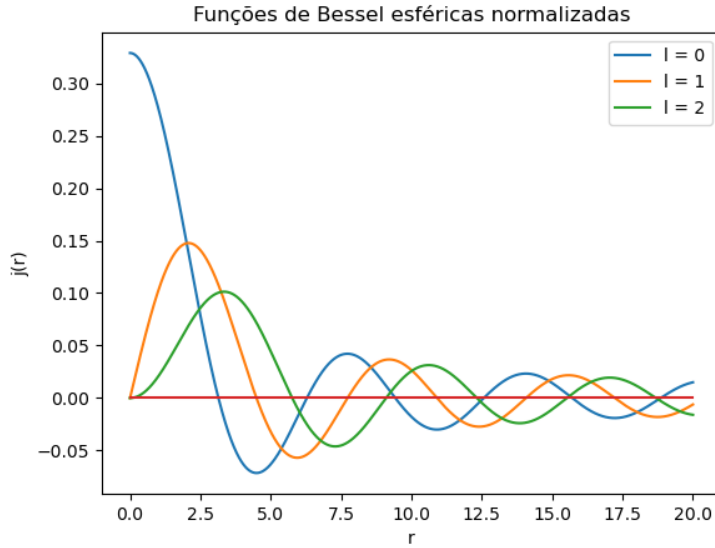
Figura 16: As primeiras funções de Bessel esféricas com $k = 1$ e $l \in \{0, 1, 2\}$ normalizadas em relação a um intervalo infinito.

A segunda solução é dada por

$$f_2(r) = r^{\alpha_2} \sum_{n=0}^{\infty} b_n r^n + g(r)$$
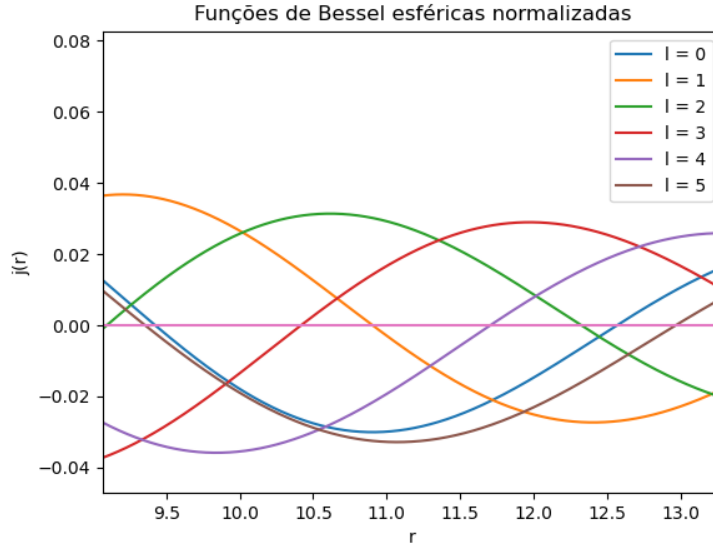
Figura 17: As primeiras seis funções de Bessel esféricas com $k = 1$ normalizadas ampliadas em um intervalo que está entre duas raízes consecutivas de j_5. Note que todas as funções de ordem l inferior a 5 possuem pelo menos uma raiz entre duas consecutivas de j_5.

onde

$$g(r) = K f_1(r) \ln |r| \quad \text{e} \quad \alpha_2 = -l(l+1)$$

Devido a aparição do logaritmo, é incerto se $f_2(r)$ converge em $r = 0$. Para descobrir, vamos usar a identidade de Abel:

Teorema 6.5 (Identidade de Abel). *Dada a equação diferencial*

$$f''(x) + p(x)f'(x) + q(x)f(x) = 0$$

O Wronskiano das duas soluções linearmente independentes é dado pela expressão

$$W(x) = C \cdot \exp\left(-\int_{x_0}^{x} p(x')\,\mathrm{d}x'\right)$$

A demonstração do teorema está seção (K) do apêndice.

Da equação (6.4) temos

$$p(r) = \frac{2}{r} \qquad q(r) = k^2 - \frac{l(l+1)}{r^2}$$

e vamos verificar o intervalo $I = (0, +\infty)$. Pelo teorema (6.5), o wronskiano em

I é dado por

$$\begin{aligned} W(r) &= C \cdot \exp\left(-\int_{r_0}^{r} \frac{2}{r'} \, \mathrm{d}r' \right) \\ &= C \cdot \exp\left(-2 \ln\left(\frac{r}{r_0} \right) \right) \\ &= C \cdot \left(\frac{r_0}{r} \right)^2 \end{aligned}$$

Isso significa que $W(r)$ diverge quando $r \to 0^+$. Mas

$$W(r) = f_1(r) f_2'(r) - f_1'(r) f_2(r)$$

e sabemos que $f_1(r)$ não diverge em $r = 0$, já que é analítica em $r = 0$. Então resta apenas que $f_2(r)$ e/ou $f_2'(r)$ divergem. Se uma função tende a $+\infty$ ou $-\infty$ em um ponto, sua derivada em módulo também tende a infinito nesse ponto e vice-versa.

Como estamos interessados apenas em funções contínuas, podemos descartar $f_2(r)$ e portanto a solução procurada é

$$R_l(r) = C \cdot f_1(r) = A \cdot \sum_{n=0}^{\infty} a_{2n} r^{2n+l}$$

Uma outra forma de encontrar as soluções da equação radial está no apêndice (H). Dentre as soluções que não divergem em $r = 0$, vamos usar a equação de Bessel esférica denotada por $j_l(kr)$.

A função de onda $\psi(r, \theta, \varphi)$ precisa ser nula em toda a superfície esférica de raio r_0, independentemente de θ ou φ. Então a parte radial precisa ser nula em r_0:

$$R(r_0) = j_l(k r_0) = 0$$

Então

$$k r_0 = \frac{\sqrt{2 m_0 E}}{\hbar} r_0 = u_{nl} \tag{376}$$

onde u_{nl} é a n-ésima raiz da função esférica de Bessel j_l. Daí temos que

$$k_{nl} = \frac{u_{nl}}{r_0}$$

Isolando o termo da energia em (376), temos que os possíveis níveis de energia são dados por

$$E_{nl} = \frac{u_{nl}^2 \hbar^2}{2 m_0 r_0^2}$$

6.4.1 Propriedades das funções de Bessel esféricas

Agora note que a equação diferencial que as funções de Bessel esféricas satisfazem,

$$r^2 R''(r) + 2r R'(r) + \left(k^2 r^2 - l(l+1) \right) R(r) = 0$$

pode ser reescrita como

$$\frac{d}{dr}\left(r^2\frac{dR}{dr}\right)+\left(k^2r^2-l(l+1)\right)R(r)=0$$

que é a equação do problema de Sturm-Liouville

$$\begin{aligned}LR+\lambda\omega(r)R&=(p(r)R'(r))'+q(r)R(r)+\lambda\omega(r)R(r)\\&=(p(r)R'(r))'+\left[q(r)+\lambda\omega(r)\right]R(r)\\&=0\end{aligned}$$

com $p(r)=r^2$, $q(r)=-l(l+1)$, $\omega(r)=r^2$, $\lambda=k^2$ e as condições de contorno

$$j_l(0)=\delta_{0l}\quad\text{e}\quad j_l(r_0)=0 \tag{377}$$

Note que usamos k^2 como o autovalor da EDO de Bessel, pois $l(l+1)$ já é o autovalor da equação diferencial de Legendre associada.

Da equação (244), o operador diferencial L será hermitiano se para quaisquer funções f e g que satisfaçam as condições de contorno (377) a relação

$$p[\overline{f'}g-\overline{f}g']\bigg|_0^{r_0}=0$$

for obedecida. Como $p(0)=0^2=0$ e $f(r_0)=g(r_0)=0$, então L é hermitiano.

Portanto as funções do conjunto $\{j_l(k_{1l}r),j_l(k_{2l}r),j_l(k_{3l}r),\ldots\}$ são autofunções de autovalor k_{nl}^2 do problema de Sturm-Liouville e formam uma base infinita ortogonal do espaço $\mathcal{L}^2(0,r_0)$ com produto interno

$$\langle f,g\rangle=\int_0^{r_0}f(r)g(r)r^2\,\mathrm{d}r$$

Além disso, o teorema (4.12) garante que

$$\int_0^{r_0}j_l\left(\frac{u_{nl}}{r_0}r\right)j_l\left(\frac{u_{ml}}{r_0}r\right)r^2\,\mathrm{d}r=C_l\delta_{nm}$$

onde C_l é uma constante que depende de l, que nesse caso é o quadrado da norma de j_l.

Além disso, satisfazem a relação de completude no espaço $\mathcal{L}^2(0,r_0)$,

$$f(r)=\sum_{n=1}^{\infty}\frac{\langle f,j_{ln}\rangle}{\langle j_{ln},j_{ln}\rangle}j_{ln}=\sum_{n=1}^{\infty}c_{ln}j_{ln}\qquad f(r)\in\mathcal{L}^2(0,r_0) \tag{378}$$

com c_{ln} sendo as coordenadas da expansão, que são dadas por

$$c_{ln}=\frac{\langle f,j_{ln}\rangle}{\langle j_{ln},j_{ln}\rangle}=\frac{\displaystyle\int_0^{r_0}f(r)j_l\left(\frac{u_{nl}}{r_0}r\right)r^2\,\mathrm{d}r}{\displaystyle\int_0^{r_0}\left|j_l\left(\frac{u_{nl}}{r_0}r\right)\right|^2r^2\,\mathrm{d}r} \tag{379}$$

e denotando

$$j_{ln} = j_l\left(\frac{u_{nl}}{r_0}r\right)$$

Note que todas as autofunções desse problema de Sturm-Liouville possuem a mesma ordem l das funções de Bessel esféricas, uma vez que os autovalores $k_{nl}^2 = (u_{nl}/r_0)^2$ estão associados às raízes de j_l, ao contrário dos polinômios de Hermite e de Legendre, cujos autovalores estão associados ao grau do polinômio.

As raízes da primeira função de Bessel esférica ($l = 0$) são múltiplos inteiros de π, pois a sua relação de recorrência com $k = 1$ é

$$a_{2n+2} = -\frac{a_{2n}}{(2n+2)\cdot(2n+3)} \quad \forall n \geq 0$$

que lembra a relação de recorrência da função $\sin(x)$,

$$b_{2n+3} = -\frac{b_{2n+1}}{(2n+2)\cdot(2n+3)} \quad \forall n \geq 0$$

porém com os coeficientes deslocados das potências ímpares para as potências pares mais baixas, no que leva a conclusão que

$$j_0(x) = \frac{\sin(x)}{x}$$

A única exceção é em $x = 0$, ponto no qual a função $j_0(x)$ converge para 1 pelo limite fundamental.

Já para as outras funções esféricas de Bessel, é necessário recorrer a recursos computacionais para calcular as suas raízes. Usamos o código em Python da seção (H.1.1) para preencher a tabela 1 com as raízes diferentes de zero de $j_l(x)$. A primeira raiz (com $l \neq 0$) sempre será $x = 0$ devido ao fator r^l do teorema de Frobenius, então não consideramos ela na tabela 1.

l	1ª raiz	2ª raiz	3ª raiz	4ª raiz	5ª raiz
0	3,1415927	6,2831853	9,4247779	12,5663706	15,7079632
1	4,4934094	7,7252518	10,9041216	14,0661939	17,2207552
2	5,7634592	9,0950113	12,3229409	15,5146030	18,6890363
3	6,9879320	10,4171185	13,6980231	16,9236212	20,1218061
4	8,1825614	11,7049071	15,0396647	18,3012559	21,5254177

Tabela 1: As cinco primeiras raízes (diferentes de zero) das funções de Bessel esférica com até sete casas decimais. Note que as raízes de j_0 são múltiplos inteiros de π.

Os exemplos seguintes possuem a intenção de ilustrar numericamente os resultados previstos pela teoria de Sturm-Liouville sobre a expansão em autofunções e, por isso, as funções não estarão necessariamente normalizadas.

Exemplo 6.1 (Partícula concentrada no centro da esfera)**.** *Queremos que a função $R(r)$ tenha uma forma parecida com a da função*

$$g(r) = \exp\left(-\left(x - \frac{1}{5}r_0\right)^2\right) \tag{380}$$

que é uma gaussiana com o pico a $1/5$ do comprimento do raio da esfera r_0. Podemos usar a fórmula da expansão em autofunções (378) *usando os coeficientes c_{ln} calculados pela expressão* (379)*.*

Usando $r_0 = 3$ unidades de comprimento, e usando as nove primeiras raízes das funções de Bessel esféricas com $l = 0$, obteremos a expansão mostrada na figura (18)*. Para isso, usamos três programas em Python: um para obter as coordenadas e as raízes das funções (código no apêndice* (H.1.1)*); outro para obter as normas $\langle j_{ln}, j_{ln}\rangle$ para um raio específico* (H.1.2)*; e um programa que calcula as coordenadas $\langle g, j_{ln}\rangle$ e imprime o gráfico da expansão em autofunções na tela* (H.1.4)*.*
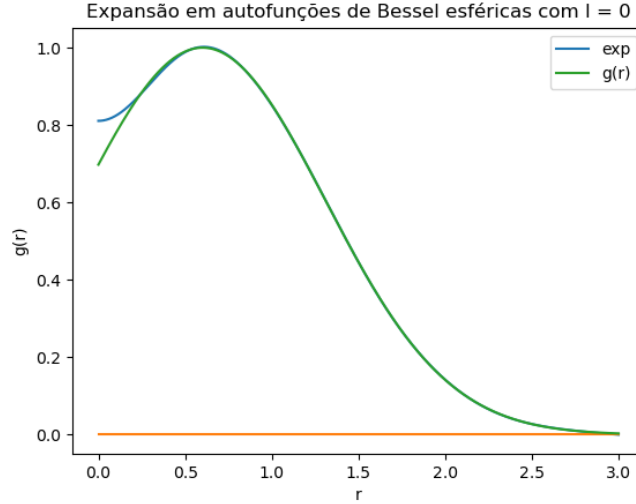
Figura 18: Expansão da função (380) nas autofunções do problema de Sturm-Liouville da equação de Bessel esférica com $l = 0$ usando as nove primeiras raízes. A linha azul é a expansão em autofunções, enquanto que a linha verde é a função (380).

Note que a convergência é muito boa em toda a esfera, exceto próximo de $r = 0$. Isso acontece porque as funções de Bessel esféricas de ordem $l = 0$ possuem a derivada de primeira ordem nula na origem, forçando a expansão a se desencontrar da função (380)*, que tem derivada de primeira ordem não nula na origem.*

Exemplo 6.2 (Partícula concentrada próximo à casca da esfera)**.** *Agora queremos que $R(r)$ tenha uma forma parecida com a da função*

$$g(r) = \exp\left(-\left(x - \frac{3}{4}r_0\right)^2\right) \tag{381}$$

que é uma gaussiana com o pico a $3/4$ do comprimento do raio da esfera r_0. Usando novamente a fórmula da expansão em autofunções (378) *com $r_0 = 3$ unidades de comprimento, e usando as nove primeiras raízes das funções de Bessel esféricas com $l = 1$, obteremos a expansão mostrada na figura* (19)*.*
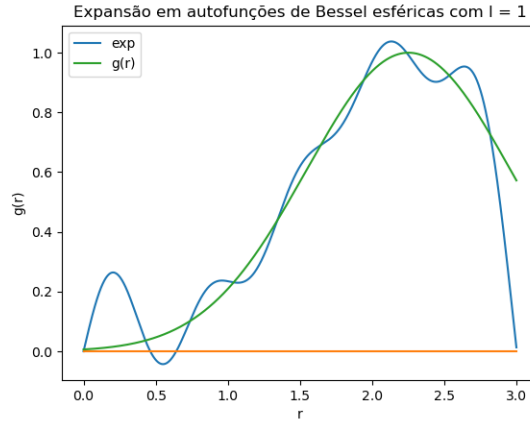
Figura 19: Expansão da função (381) nas autofunções do problema de Sturm-Liouville da equação de Bessel esférica com $l = 1$ usando as nove primeiras raízes. A linha azul é a expansão em autofunções, enquanto que a linha verde é a função (381).

Note que a expansão obrigatoriamente será nula no raio da esfera, uma vez que todas as funções do conjunto $\{j_l(k_{11}r), j_l(k_{21}r), j_l(k_{31}r), \ldots\}$ possuem raízes em $r = r_0$, enquanto que a gaussiana não é nula nesse ponto. Isso provoca uma mudança abrupta da expansão em autofunções próximo ao raio da esfera. A convergência pode ser melhorada aumentando o número de raízes utilizadas, entretanto isso não foi possível devido a divergência da série finita de Taylor das funções de Bessel esféricas, o que distorce os valores das raízes quando r é muito grande.

6.5 Harmônicos esféricos

Devido à frequência com que a equação das funções associadas de Legendre e a equação azimutal aparecem em vários problemas da física (como potenciais elétricos, ondas esféricas e aplicações da mecânica quântica) vamos denominar

o produto de $\Theta(\theta)$ com $\Phi(\phi)$ como Y_{lm}:

$$\begin{aligned} Y_{lm}(\theta,\varphi) &= \Theta(\theta)\Phi(\varphi) \\ &= C_{lm}\overline{P}_l^m(\cos(\theta))e^{im\varphi} \end{aligned}$$

que satisfaz a equação diferencial

$$\frac{1}{\sin(\theta)}\left[\frac{\partial}{\partial\theta}\left(\sin(\theta)\frac{\partial}{\partial\theta}Y_{lm}(\theta,\varphi)\right)\right] - \frac{m^2}{\sin^2(\theta)}Y_{lm}(\theta,\varphi) + \left(\frac{\partial^2}{\partial\varphi^2}+m^2\right)Y_{lm}(\theta,\varphi) = -l(l+1)Y_{lm}(\theta,\varphi) \quad (382)$$

uma vez que

$$\frac{1}{\sin(\theta)}\left[\frac{d}{d\theta}\left(\sin(\theta)\Theta'(\theta)\right)\right] - \frac{m^2}{\sin^2(\theta)}\Theta(\theta) = -l(l+1)\Theta(\theta)$$

e

$$\Phi''(\varphi) + m^2\Phi(\varphi) = 0$$

As funções $Y_{lm}(\theta,\varphi)$, com $|m| \leq l$, são chamadas de harmônicos esféricos. Como essas funções aparecem multiplicando soluções de outras equações diferenciais ordinárias, é conveniente normalizá-las em θ e φ para que a normalização da solução final só dependa das outras variáveis das outras soluções. Para isso, os coeficientes C_{lm} são escolhidos de forma que

$$\oiint_S |Y_{lm}(\theta,\varphi)|^2 \sin(\theta)\,\mathrm{d}\theta\,\mathrm{d}\varphi = 1$$

Dado os índices m, m', l, l', temos que

$$\oiint_S \overline{Y_{lm}(\theta,\varphi)}Y_{l'm'}(\theta,\varphi)\sin(\theta)\,\mathrm{d}\theta\,\mathrm{d}\varphi = \overline{C_{lm}}C_{l'm'}\int_0^\pi \overline{P}_l^m(\cos(\theta))\overline{P}_{l'}^{m'}(\cos(\theta))\sin(\theta)\,\mathrm{d}\theta\int_0^{2\pi} e^{-im\varphi}e^{im'\varphi}\,\mathrm{d}\varphi$$

Vamos considerar $m \neq m'$. Então

$$\begin{aligned} \int_0^{2\pi} e^{-im\varphi}e^{im'\varphi}\,\mathrm{d}\varphi &= \int_0^{2\pi} e^{i(m'-m)\varphi}\,\mathrm{d}\varphi \\ &= \int_0^{2\pi} \cos((m'-m)\varphi) + i\sin((m'-m)\varphi)\,\mathrm{d}\varphi \\ &= \left.\frac{\sin((m'-m)\varphi)}{m'-m}\right|_0^{2\pi} - i\left.\frac{\cos((m'-m)\varphi)}{m'-m}\right|_0^{2\pi} \\ &= 0 \end{aligned}$$

Por outro lado, se $m' = m$,

$$\int_0^{2\pi} e^{-im\varphi} e^{im'\varphi}\, \mathrm{d}\varphi = \int_0^{2\pi} \mathrm{d}\varphi = 2\pi$$

ou seja,

$$\int_0^{2\pi} e^{-im\varphi} e^{im'\varphi}\, \mathrm{d}\varphi = 2\pi\delta_{mm'}$$

Agora considerando $m' = m$ e $x = \cos(\theta)$, temos

$$\begin{aligned}\int_0^{\pi} \overline{P}_l^m(\cos(\theta))\overline{P}_{l'}^{m'}(\cos(\theta))\sin(\theta)\,\mathrm{d}\theta &= \int_{-1}^{1} \overline{P}_l^m(x)\overline{P}_{l'}^m(x)\,\mathrm{d}x \\ &= \delta_{ll'}\end{aligned}$$

e portanto

$$\oiint_S \overline{Y_{lm}(\theta,\varphi)} Y_{l'm'}(\theta,\varphi)\sin(\theta)\,\mathrm{d}\theta\,\mathrm{d}\varphi = \overline{C_{lm}} C_{lm} 2\pi\delta_{mm'}\delta_{ll'}$$

Para que Y_{lm} seja normalizado, devemos escolher $C_{lm} = \dfrac{1}{\sqrt{2\pi}}$ e então

$$\oiint_S \overline{Y_{lm}(\theta,\varphi)} Y_{l'm'}(\theta,\varphi)\sin(\theta)\,\mathrm{d}\theta\,\mathrm{d}\varphi = \delta_{ll'}\delta_{mm'}$$

As projeções cilíndricas de alguns dos harmônicos esféricos (semelhante à projeção de Mercator) estão ilustradas nas figuras (20),(21) e (22). Usamos apenas a parte real da função azimutal $\Phi(\varphi)$ para ilustrar mudanças em φ, pois $|\Phi(\varphi)|$ e $\overline{\Phi(\varphi)}\Phi(\varphi)$ são sempre iguais a 1. Os gráficos foram construídos usando o código Python da seção (I.2.4), que usam os coeficientes e as normas calculadas das funções associadas de Legendre pelos programas (I.2.1) e (I.2.2).

A solução final para o problema da partícula confinada em uma esfera é a solução geral da equação de Schrödinger independente do tempo multiplicada pelo fator $\exp\left(-\frac{iE_{nl}}{\hbar}t\right)$:

$$\begin{aligned}\psi_{lmn}(r,\theta,\varphi,t) &= A_{nl}\cdot j_l(kr)Y_{lm}(\theta,\varphi)\exp\left(-\frac{iE_{nl}}{\hbar}t\right) \\ &= A_{nl}\cdot j_l\left(\frac{\sqrt{2m_0E_{nl}}}{\hbar}r\right)Y_{lm}(\theta,\varphi)\exp\left(-\frac{iE_{nl}}{\hbar}t\right) \\ &= A_{nl}\cdot j_l\left(\frac{u_{nl}}{r_0}r\right)Y_{lm}(\theta,\varphi)\exp\left(-\frac{iE_{nl}}{\hbar}t\right)\end{aligned}$$

onde A_{nl} é uma constante tal que

$$A_{nl}\int_0^{r_0} j_l\left(\frac{u_{nl}}{r_0}r\right)\mathrm{d}r = 1$$

para que a função de onda seja normalizada.

Em geral, temos

$$\psi(r,\theta,\varphi,t) = \sum_{n=1}^{\infty}\sum_{l=0}^{\infty}\sum_{m=0}^{l} A_{nl}\cdot j_l\left(\frac{u_{nl}}{r_0}r\right)Y_{lm}(\theta,\varphi)\exp\left(-\frac{iE_{nl}}{\hbar}t\right)$$
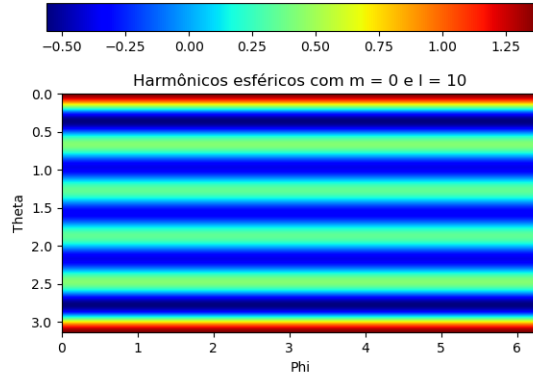
Figura 20: Projeção cilíndrica do harmônico esférico com $m = 0$ e $l = 10$.
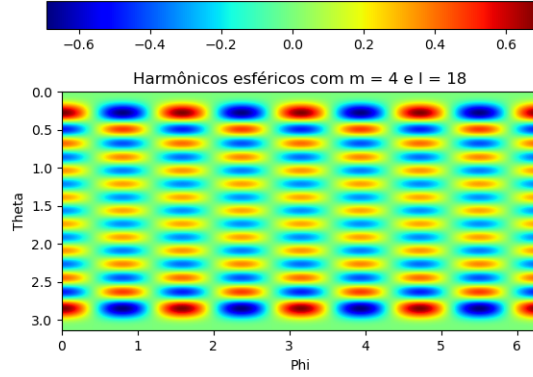
Figura 21: Projeção cilíndrica do harmônico esférico com $m = 4$ e $l = 18$.

Conforme vimos na seção (6.3.1), se $m \geq 1$ então as funções de Legendre associadas P_l^m possuirão $2 + l - m$ raízes no domínio trabalhado. Isso significa que quanto maior for a diferença $l - m$, maiores serão as lacunas no espaço (em
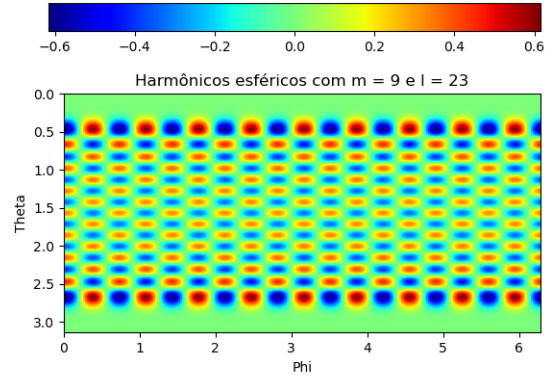
Figura 22: Projeção cilíndrica do harmônico esférico com $m = 9$ e $l = 23$.

relação ao ângulo polar θ) onde a partícula terá chance nula de ser detectada e a densidade de probabilidade ficará mais concentrada em regiões específicas.

7 O átomo de hidrogênio

Nesta seção, trataremos o problema de descobrir a função de onda que rege um átomo de hidrogênio (eletricamente neutro e com apenas um próton no núcleo) assim como os níveis de energia que o elétron pode assumir enquanto preso ao núcleo.

Na equação de Schrödinger independente do tempo,

$$-\frac{\hbar^2}{2m_0}\nabla^2\Psi + [V(r) - E]\Psi = 0 \tag{383}$$

usaremos o potencial de Coulomb,

$$V(r) = -\frac{e^2}{4\pi\epsilon_0}\frac{1}{r} \tag{384}$$

com $E < 0$ para representar o estado ligado do elétron.

A equação (383) pode ser reescrita como

$$\nabla^2\Psi + \frac{2m_0}{\hbar^2}[E - V(r)]\Psi = 0 \tag{385}$$

Como o potencial só depende de r, então a equação diferencial para θ e φ será idêntica à (382) (harmônicos esféricos), como veremos a seguir.

Abrindo o operador nabla em (385), temos

$$\frac{2}{r}\frac{\partial\psi}{\partial r} + \frac{\partial^2\psi}{\partial r^2} + \frac{\cot(\theta)}{r^2}\frac{\partial\psi}{\partial\theta} + \frac{1}{r^2}\frac{\partial^2\psi}{\partial\theta^2} + \frac{1}{r^2\sin^2(\theta)}\frac{\partial^2\psi}{\partial\varphi^2} + \frac{2m_0}{\hbar^2}(E - V(r))\psi = 0$$

Multiplicando ambos os lados por r^2, obtemos

$$2r\frac{\partial\psi}{\partial r} + r^2\frac{\partial^2\psi}{\partial r^2} + \cot(\theta)\frac{\partial\psi}{\partial\theta} + \frac{\partial^2\psi}{\partial\theta^2} + \frac{1}{\sin^2(\theta)}\frac{\partial^2\psi}{\partial\varphi^2} + \frac{2m_0r^2}{\hbar^2}(E - V(r))\psi = 0$$

Isolando os termos com r,

$$r^2\frac{\partial^2\psi}{\partial r^2} + 2r\frac{\partial\psi}{\partial r} + \frac{2m_0r^2}{\hbar^2}(E - V(r))\psi = -\left[\cot(\theta)\frac{\partial\psi}{\partial\theta} + \frac{\partial^2\psi}{\partial\theta^2} + \frac{1}{\sin^2(\theta)}\frac{\partial^2\psi}{\partial\varphi^2}\right]$$

Repetindo o procedimento da separação de variáveis, onde $\psi(r,\theta,\varphi) = R(r)\Theta(\theta)\Phi(\varphi)$, temos que

$$\frac{r^2}{R(r)}\frac{d^2R}{dr^2} + \frac{2r}{R(r)}\frac{dR}{dr} + \frac{2m_0r^2}{\hbar^2}(E - V(r)) = -\left[\frac{\cot(\theta)}{Y}\frac{\partial Y}{\partial\theta} + \frac{1}{Y}\frac{\partial^2 Y}{\partial\theta^2} + \frac{1}{Y\sin^2(\theta)}\frac{\partial^2 Y}{\partial\varphi^2}\right]$$

com $Y = \Theta(\theta)\Phi(\varphi)$. Como r, θ e φ são variáveis independentes, a equação só pode ser verdade se

$$-\left[\frac{\cot(\theta)}{Y}\frac{\partial Y}{\partial\theta} + \frac{1}{Y}\frac{\partial^2 Y}{\partial\theta^2} + \frac{1}{Y\sin^2(\theta)}\frac{\partial^2 Y}{\partial\varphi^2}\right] = \lambda_2$$

e

$$\frac{r^2}{R(r)}\frac{d^2R}{dr^2} + \frac{2r}{R(r)}\frac{dR}{dr} + \frac{2m_0r^2}{\hbar^2}(E - V(r)) = \lambda_2$$

para alguma constante λ_2. Segue daí que

$$r^2\frac{d^2R}{dr^2} + 2r\frac{dR}{dr} + \left[\frac{2m_0r^2}{\hbar^2}(E - V(r)) - \lambda_2\right]R(r) = 0$$

e

$$\frac{\cot(\theta)}{Y}\frac{\partial Y}{\partial\theta} + \frac{1}{Y}\frac{\partial^2 Y}{\partial\theta^2} + \frac{1}{Y\sin^2(\theta)}\frac{\partial^2 Y}{\partial\varphi^2} = -\lambda_2 \tag{386}$$

Multiplicando a equação (386) por $\sin^2(\theta)$, e separando os termos com θ e φ,

$$\frac{1}{Y}\sin(\theta)\cos(\theta)\frac{\partial Y}{\partial\theta} + \frac{\sin^2(\theta)}{Y}\frac{\partial^2 Y}{\partial\theta^2} + \lambda_2\sin^2(\theta) = -\frac{1}{Y}\frac{\partial^2 Y}{\partial\varphi^2} = -\lambda_1$$

onde o primeiro membro só depende de θ e o segundo de φ. Então temos duas equações,

$$\frac{\partial^2 Y}{\partial\varphi^2} + \lambda_1 Y = 0 \tag{387}$$

$$\sin(\theta)\cos(\theta)\frac{\partial Y}{\partial\theta} + \sin^2(\theta)\frac{\partial^2 Y}{\partial\theta^2} - \lambda_1 Y = -\lambda_2\sin^2(\theta)Y \tag{388}$$

Dividindo (388) por $\sin^2(\theta)$, ficamos com

$$\cot(\theta)\frac{\partial Y}{\partial\theta} + \frac{\partial^2 Y}{\partial\theta^2} - \frac{\lambda_1}{\sin^2(\theta)}Y = -\lambda_2 Y$$

Notando novamente que

$$\begin{aligned}\cot(\theta)\frac{\partial Y}{\partial\theta} + \frac{\partial^2 Y}{\partial\theta^2} &= \frac{1}{\sin(\theta)}\cdot\left[\cos(\theta)\frac{\partial Y}{\partial\theta} + \sin(\theta)\frac{\partial^2 Y}{\partial\theta^2}\right] \\ &= \frac{1}{\sin(\theta)}\cdot\left[\frac{\partial}{\partial\theta}\left(\sin(\theta)\frac{\partial Y}{\partial\theta}\right)\right]\end{aligned}$$

temos

$$\frac{1}{\sin(\theta)}\cdot\left[\frac{\partial}{\partial\theta}\left(\sin(\theta)\frac{\partial Y}{\partial\theta}\right)\right] - \frac{\lambda_1}{\sin^2(\theta)}Y = -\lambda_2 Y \tag{389}$$

Somando a equação (387) com (389), obtemos

$$\frac{1}{\sin(\theta)}\cdot\left[\frac{\partial}{\partial\theta}\left(\sin(\theta)\frac{\partial}{\partial\theta}Y(\theta,\varphi)\right)\right] - \frac{\lambda_1}{\sin^2(\theta)}Y(\theta,\varphi) + \left(\frac{\partial^2}{\partial\varphi^2} + \lambda_1\right)Y(\theta,\varphi) = -\lambda_2 Y(\theta,\varphi) \tag{390}$$

Agora compare (390) com a equação dos harmônicos esféricos (382), que é

$$\frac{1}{\sin(\theta)}\left[\frac{\partial}{\partial\theta}\left(\sin(\theta)\frac{\partial}{\partial\theta}Y_{lm}(\theta,\varphi)\right)\right]-\frac{m^2}{\sin^2(\theta)}Y_{lm}(\theta,\varphi)+$$
$$\left(\frac{\partial^2}{\partial\varphi^2}+m^2\right)Y_{lm}(\theta,\varphi)=-l(l+1)Y_{lm}(\theta,\varphi)$$

Conforme vimos no problema do potencial esfericamente simétrico, as únicas soluções não-nulas para $Y(\theta,\varphi)$ que são contínuas no intervalo $\theta\in[0,\pi]$ e satisfazem a condição de contorno periódica $Y(\theta,\varphi)=Y(\theta,\varphi+2\pi)$ para qualquer φ são as que

$$\lambda_1=m^2 \quad \text{e} \quad \lambda_2=l(l+1)$$

com l e m inteiros, obedecendo $|m|\leq l$. Portanto, as soluções da equação (390) são as funções $Y_{lm}(\theta,\varphi)$ que vimos na seção (6.5).

7.1 A equação radial

Já para a equação para a coordenada r, temos

$$r^2\frac{d^2R}{dr^2}+2r\frac{dR}{dr}+\left[\frac{2m_0r^2}{\hbar^2}(E-V(r))-l(l+1)\right]R(r)=0 \tag{391}$$

Substituindo $V(r)$ pelo potencial do átomo de hidrogênio (384) em (391), obtemos

$$r^2\frac{d^2R}{dr^2}+2r\frac{dR}{dr}+\left[\frac{2m_0r^2}{\hbar^2}\left(E+\frac{e^2}{4\pi\epsilon_0}\frac{1}{r}\right)-l(l+1)\right]R(r)=0$$

ou

$$r^2\frac{d^2R}{dr^2}+2r\frac{dR}{dr}+\left[\frac{2m_0E}{\hbar^2}r^2+\frac{2m_0e^2}{4\pi\epsilon_0\hbar^2}r-l(l+1)\right]R(r)=0 \tag{392}$$

Para tornar a equação adimensional, faremos a substituição

$$\frac{2m_0E}{\hbar^2}r^2=-x^2$$

onde o sinal de menos foi escolhido por causa da energia E negativa.

Então

$$x=\frac{\sqrt{-2m_0E}}{\hbar}r$$

de modo que

$$\frac{2m_0E}{\hbar^2}r^2=-x^2$$

e

$$\begin{aligned}\frac{2m_0e^2}{4\pi\epsilon_0\hbar^2}r &= \frac{2m_0e^2}{4\pi\epsilon_0\hbar^2}\cdot\frac{\hbar x}{\sqrt{-2m_0E}}\\ &= \frac{2m_0e^2}{4\pi\epsilon_0\hbar\sqrt{-2m_0E}}\cdot x\\ &= x_0x\end{aligned}$$

onde

$$x_0 = \frac{2m_0e^2}{4\pi\epsilon_0\hbar\sqrt{-2m_0E}}$$

Se denotarmos $\dfrac{\sqrt{-2m_0E}}{\hbar} = k$, $E \neq 0$, de modo que $x = kr$ e $r = \frac{x}{k}$, então

$$\frac{d}{dr} = \frac{dx}{dr}\frac{d}{dx} = k\frac{d}{dx}$$

$$\frac{d^2}{dr^2} = \frac{d}{dr}\frac{d}{dr} = k^2\frac{d^2}{dx^2}$$

Substituindo essas expressões na equação diferencial (392), obtemos

$$\frac{x^2}{k^2}k^2\frac{d^2R}{dx^2} + 2\frac{x}{k}k\frac{dR}{dx} + \left[-x^2 + x_0x - l(l+1)\right]R(x) = 0$$

$$x^2\frac{d^2R}{dx^2} + 2x\frac{dR}{dx} - x^2\left[1 - \frac{x_0}{x} + \frac{l(l+1)}{x^2}\right]R(x) = 0 \qquad (393)$$

A equação (393) é a que precisamos resolver para encontrar a parte radial da solução da função de onda. Iremos resolver por partes.

Quando x é muito grande, os termos com $1/x$ e $1/x^2$ serão desprezíveis, restando apenas

$$x^2\frac{d^2R}{dx^2} + 2x\frac{dR}{dx} - x^2R(x) = 0 \qquad (394)$$

Sabendo que

$$x^2\frac{d^2R}{dx^2} + 2x\frac{dR}{dx} = \frac{d}{dx}\left(x^2\frac{dR}{dx}\right)$$

então (394) resulta na equação

$$\frac{d}{dx}\left(x^2\frac{dR}{dx}\right) = x^2R(x) \qquad (395)$$

A expressão (395) nos motiva a fazer a substituição $u(x) = x \cdot R(x)$ e então

$$u'(x) = xR'(x) + R(x) \qquad u''(x) = xR''(x) + 2R'(x)$$

Agora note que se multiplicarmos $u''(x)$ por x, obteremos

$$x \cdot u''(x) = x^2\frac{d^2R}{dx^2} + 2x\frac{dR}{dx}$$

que são justamente os dois primeiros termos da equação (394), sendo que o terceiro pode ser escrito como $x^2R(x) = x \cdot u(x)$. Portanto a equação diferencial se torna

$$xu''(x) - x \cdot u(x) = 0$$

$$u''(x) = u(x)$$

A solução dessa equação diferencial é a trivial combinação linear de exponenciais reais,

$$u(x) = c_1 e^x + c_2 e^{-x}$$

o que significa

$$R(x) = c_1 \frac{e^x}{x} + c_2 \frac{e^{-x}}{x}$$

Queremos que $\psi(r(x), \theta, \varphi) = R(x) Y_{lm}(\theta, \varphi)$ seja quadrado integrável em todas as coordenadas, e para isso $R(x)$ precisa ser quadrado integrável com respeito a x, que varia de 0 a $+\infty$ (uma vez que x é r a menos de um fator constante positivo). Das duas soluções linearmente independentes de $R(x)$ para x muito grande, precisamos, portanto, descartar o termo e^x/x, uma vez que diverge em $+\infty$, mesmo com x no denominador.

Retornando à EDO (393), temos que o comportamento assintótico de $R(x)$ quando $x \to +\infty$ deve ser aproximado por

$$R(x) \approx c_2 \frac{e^{-x}}{x}$$

Quando x, no intervalo $0 < x < 1$, é muito próximo de zero, o fator $\frac{l(l+1)}{x^2}$ predomina, de modo que a equação diferencial possa ser aproximada por

$$x^2 \frac{d^2 R}{dx^2} + 2x \frac{dR}{dx} - l(l+1) R(x) = 0 \tag{396}$$

que é uma equação de Euler-Cauchy. Fazendo a substituição $R(x) = x^v$, temos que

$$R'(x) = v x^{v-1} \qquad R''(x) = v(v-1) x^{v-2}$$

Substituindo em (396), obtemos

$$v(v-1) x^v + 2v x^v - l(l+1) x^v = 0$$

Queremos que a equação seja válida para qualquer x, em especial $x \neq 0$. Dividindo ambos os lados por x^v, temos que

$$\begin{aligned} 0 &= v(v-1) + 2v - l(l+1) \\ &= v^2 + v - l(l+1) \\ &= v(v+1) - l(l+1) \end{aligned}$$

As soluções para v são as mesmas da equação indicial (368) que resolvemos na EDO de Bessel esférica:

$$v_1 = l \quad \text{e} \quad v_2 = -(l+1)$$

Isso significa que a solução da EDO (393) tende à função

$$R(x) \approx d_1 x^l + \frac{d_2}{x^{l+1}} \tag{397}$$

quando $x \to 0^+$.

A condição de que $R(x)$ seja quadrado integrável em $[0, +\infty)$ implica que $R(x)$ não pode divergir em $x = 0$. Consequentemente, vamos desconsiderar o segundo termo da solução (397) de forma que

$$x \to 0^+ \implies R(x) \approx d_1 x^l$$

Então a solução da EDO (393),

$$x^2 \frac{d^2R}{dx^2} + 2x\frac{dR}{dx} - x^2\left[1 - \frac{x_0}{x} + \frac{l(l+1)}{x^2}\right] R(x) = 0 \tag{398}$$

deve ter alguma forma do tipo

$$R(x) = x^l e^{-x} u(x)$$

para alguma função desconhecida $u(x)$. Fazendo essa substituição, na primeira derivada temos

$$R'(x) = x^{l-1}e^{-x}(l-x)u(x) + x^l e^{-x}u'(x)$$

e então

$$2xR'(x) = 2x^l e^{-x}(l-x)u(x) + 2x^{l+1}e^{-x}u'(x)$$

Para a segunda derivada, temos

$$R''(x) = l(l-1)x^{l-2}e^{-x}u(x) + x^l e^{-x}u(x) + x^l e^{-x}u''(x) + \\ 2lx^{l-1}e^{-x}u'(x) - 2lx^{l-1}e^{-x}u(x) - 2x^l e^{-x}u'(x)$$

A última expressão pode ser simplificada para

$$R''(x) = x^l e^{-x}u''(x) + 2x^{l-1}e^{-x}\bigg[l - x\bigg]u'(x) + \\ + x^{l-2}e^{-x}\bigg[l(l-1) + x^2 - 2lx\bigg]u(x)$$

Logo

$$x^2R''(x) = x^{l+2}e^{-x}u''(x) + 2x^{l+1}e^{-x}\bigg[l - x\bigg]u'(x) + \\ + x^l e^{-x}\bigg[l(l-1) + x^2 - 2lx\bigg]u(x)$$

Agora vamos substituir essas expressões na equação diferencial (398), que é equivalente à EDO

$$x^2\frac{d^2R}{dx^2} + 2x\frac{dR}{dx} - x^2R(x) + x_0xR(x) - l(l+1)R(x) = 0 \tag{399}$$

Portanto a EDO em termos de $u(x)$ é

$$x^{l+2}e^{-x}u''(x) + 2x^{l+1}e^{-x}\big[l+1-x\big]u'(x) + \\ x^l e^{-x}\Bigg[2(l-x) + l(l-1) - 2lx + x^2 - x^2 + x_0 x - l(l+1)\Bigg]u(x) = 0$$

No fator entre parênteses de $u(x)$, o termo quadrático x^2 desaparece, assim como os termos $2l$, $l(l-1)$ e $-l(l+1)$, e consequentemente

$$x^{l+2}e^{-x}u''(x) + 2x^{l+1}e^{-x}\big[l+1-x\big]u'(x) + \\ x^l e^{-x}\Bigg[(x_0 - 2(1+l))\cdot x\Bigg]u(x) = 0$$

Por fim, podemos dividir ambos os lados por $x^{l+1}e^{-x}$ (pois a equação é válida para qualquer $x \neq 0$) para obtermos a equação diferencial

$$xu''(x) + 2\big[l+1-x\big]u'(x) + \big[x_0 - 2(1+l)\big]u(x) = 0 \tag{400}$$

A EDO (400) é uma variante da equação de Laguerre associada, que é

$$xv''(x) + \big[l+1-x\big]v'(x) + \big[x_0 - 2(1+l)\big]v(x) = 0$$

A diferença entre as duas equações é somente um fator 2 em $u'(x)$. Entretanto ainda é possível estabelecer uma relação entre a solução $u(x)$ da nossa equação e a solução $v(x)$ da equação de Laguerre associada, como veremos adiante.

Devido ao ponto singular em $x_0 = 0$, devemos usar o método de Frobenius. Na forma

$$u''(x) + \frac{2\big[l+1-x\big]}{x}u'(x) + \frac{\big[x_0 - 2(1+l)\big]}{x}u(x) = 0 \tag{401}$$

temos

$$P(x) = \frac{2\big[l+1-x\big]}{x} \quad \text{e} \quad Q(x) = \frac{\big[x_0 - 2(1+l)\big]}{x}$$

Vemos que $x_0 = 0$ é um ponto singular regular, pois

$$\lim_{x\to 0} x\cdot P(x) = \lim_{x\to 0} 2\big[l+1-x\big] = 2(l+1) \in \mathbb{R}$$

e

$$\lim_{x\to 0} x^2\cdot Q(x) = \lim_{x\to 0} x\big[x_0 - 2(1+l)\big] = 0 \in \mathbb{R}$$

A equação indicial é

$$\alpha(\alpha-1) + 2(l+1)\alpha = 0$$

A primeira solução é $\alpha_1 = 0$, e para a segunda,

$$\alpha_2 - 1 + (2l+1) = 0 \implies \alpha_2 = -2l$$

Como $\alpha_1 - \alpha_2$ é um inteiro não-negativo, pelo caso inteiro do teorema de Frobenius (6.1), duas soluções linearmente independentes são

$$u_1(x) = r^{\alpha_1} \sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} a_n x^n \tag{402}$$

e

$$\begin{aligned} u_2(x) &= r^{\alpha_2} \sum_{n=0}^{\infty} b_n x^n + \big[\delta_{0l}(1-K) + K\big] u_1(x) \ln|x| \\ &= \frac{1}{r^{2l}} \sum_{n=0}^{\infty} b_n x^n + \big[\delta_{0l}(1-K) + K\big] u_1(x) \ln|x| \end{aligned}$$

com K arbitrário. O termo $\delta_{0l}(1-K)+K$ é justificado pela exigência do teorema de Frobenius (6.1) de que seja igual a 1 caso $l = 0$.

Para a primeira solução, temos

$$u_1'(x) = \sum_{n=0}^{\infty} n a_n x^{n-1} = \sum_{n=0}^{\infty} (n+1) a_{n+1} x^n$$

$$u_1''(x) = \sum_{n=0}^{\infty} n(n-1) a_n x^{n-2} = \sum_{n=0}^{\infty} (n+1) n a_{n+1} x^{n-1}$$

Substituindo essas expressões na equação (400), obtemos

$$x \sum_{n=0}^{\infty} n(n+1) a_{n+1} x^{n-1} + 2(l+1) \sum_{n=0}^{\infty} (n+1) a_{n+1} x^n - \\ 2x \sum_{n=0}^{\infty} n a_n x^{n-1} + \big[x_0 - 2(1+l)\big] \sum_{n=0}^{\infty} a_n x^n = 0$$

$$\sum_{n=0}^{\infty} \Bigg[(n(n+1) + 2(l+1)(n+1)) a_{n+1} + (x_0 - 2(1+l) - 2n) a_n \Bigg] x^n = 0$$

Devido a independência linear dos monômios x^n, temos que

$$(n(n+1) + 2(l+1)(n+1)) a_{n+1} + (x_0 - 2(1+l) - 2n) a_n = 0 \quad \forall n \geq 0$$

e portanto a relação de recorrência é

$$a_{n+1} = \frac{2n + 2(1+l) - x_0}{n(n+1) + 2(l+1)(n+1)} a_n \quad \forall n \geq 0$$

Fatorando o numerador e denominador, temos a sua versão mais simplificada,

$$a_{n+1} = \frac{2(1+l+n) - x_0}{(n+1)(n+2l+2)} a_n \quad \forall n \geq 0 \tag{403}$$

com $a_0 \neq 0$ arbitrário.

Queremos encontrar o comportamento da recorrência (403) quando n é muito grande (e consequentemente quando x é muito grande). Uma possível aproximação é dada pela expressão

$$a_{n+1} \approx \frac{2n}{n^2} a_n = \frac{2}{n} a_n$$

Entretanto ela é problemática para obter a_1 usando a_0 (e consequentemente para obter os demais coeficientes a_n também). Então vamos usar uma aproximação melhor:

$$a_{n+1} \approx \frac{2n}{(n+1)n} a_n = \frac{2}{n+1} a_n$$

Como é uma relação simples, vamos aplicá-la recursivamente:

$$a_{n+1} \approx \frac{2}{n+1} \cdot \frac{2}{n} \cdot \frac{2}{n-1} \cdot \frac{2}{n-2} \cdot \ldots \cdot \frac{2}{1} a_0$$

É evidente que a constante 2 no numerador resultará em uma potência 2^{n+1} multiplicando a_0, enquanto que o denominador será o fatorial de $n+1$:

$$a_{n+1} \approx \frac{2^{n+1}}{(n+1)!} a_0$$

ou seja,

$$a_n = \frac{2^n}{n!} a_0$$

Mas esses são justamente os coeficientes da série de Taylor da função $a_0 e^{2x}$:

$$\sum_{n=0}^{\infty} a_0 \frac{2^n}{n!} x^n = \sum_{n=0}^{\infty} \frac{a_0}{n!} (2x)^n = a_0 e^{2x}$$

Isso significa que se x for muito grande, a solução da equação diferencial

$$x^2 \frac{d^2 R}{dx^2} + 2x \frac{dR}{dx} - x^2 \left[1 - \frac{x_0}{x} + \frac{l(l+1)}{x^2} \right] R(x) = 0 \tag{404}$$

se comportará como

$$u(x) \approx x^l e^{-x} e^{2x} = x^l e^x$$

que diverge quando $x \to \infty$ e portanto essa solução não será quadrado-integrável.

Revendo atentamente a relação de recorrência

$$a_{n+1} = \frac{2(1+l+n) - x_0}{(n+1)(n+2l+2)} a_n \quad \forall n \geq 0 \tag{405}$$

podemos ver que se x_0 for um número tal que $2(1+l+n) - x_0 = 0$ para algum $n = n'$, então $a_{n+1} = 0$. Isso trunca a série (402) de $u_1(x)$, resultando em um polinômio $U_n^l(x)$. Ou seja, a solução

$$u(x) = x^l e^{-x} U_n^l(x)$$

será quadrado integrável e, portanto, a solução para $\psi(\vec{\mathbf{r}}, t)$ será normalizável. Usando a relação de recorrência (405) com

$$x_0 = 2(1 + l + n')$$

resulta em

$$a_{n+1} = 2 \cdot \frac{n - n'}{(n+1)(n+2l+2)} a_n \tag{406}$$

Os primeiros termos dessa relação de recorrência são

$$a_1 = a_{0+1} = \frac{(-1)^1 \cdot 2 \cdot n'}{2l+2} a_0 \qquad a_2 = a_{1+1} = \frac{2^2 \cdot (-1)^2 \cdot n'(n'-1)}{2 \cdot (2l+2)(2l+3)} a_0$$

$$a_3 = a_{2+1} = \frac{2^3 \cdot (-1)^3 \cdot n'(n'-1)(n'-2)}{3 \cdot 2 \cdot (2l+2)(2l+3)(2l+4)} a_0$$

$$a_4 = a_{3+1} = \frac{2^4 \cdot (-1)^4 \cdot n'(n'-1)(n'-2)(n'-3)}{4 \cdot 3 \cdot 2 \cdot (2l+2)(2l+3)(2l+4)(2l+5)} a_0$$

Vendo esse padrão, é plausível tentar

$$a_n = a_0 \frac{2^n \cdot (-1)^n \cdot (n')!}{n! \cdot (n'-n)!} \cdot \prod_{k=2}^{n+1} \frac{1}{2l+k} \qquad \forall n,\ 1 \leq n \leq n' \tag{407}$$

Colocando a expressão (407) na relação de recorrência (406), temos que

$$\begin{aligned}
a_{n+1} &= 2 \cdot a_0 \frac{n - n'}{(n+1)(n+2l+2)} a_n \\
&= 2 \cdot a_0 \frac{(-1) \cdot (n'-n)}{(n+1)(n+2l+2)} \frac{2^n \cdot (-1)^n \cdot (n')!}{n! \cdot (n'-n)!} \cdot \prod_{k=2}^{n+1} \frac{1}{2l+k} \\
&= a_0 \frac{2^{n+1} \cdot (-1)^{n+1} \cdot (n')!}{(n+1)! \cdot (n'-(n+1))!} \cdot \prod_{k=2}^{n+2} \frac{1}{2l+k}
\end{aligned}$$

que é (407) substituindo n por $n+1$. Portanto a relação foi provada.

Podemos simplificar a expressão (407) para

$$a_n = (-2)^n a_0 \binom{n'}{n} \cdot \prod_{k=2}^{n+1} \frac{1}{2l+k} \qquad \forall n,\ 1 \leq n \leq n'$$

onde

$$\binom{n'}{n} = \frac{(n')!}{n!(n'-n)!}$$

é o coeficiente binomial. Além disso, como o produtório age n vezes, podemos incluir o fator $(-2)^n$ nele:

$$a_n = a_0 \binom{n'}{n} \cdot \prod_{k=2}^{n+1} \left(-\frac{2}{2l+k} \right) \qquad \forall n,\ 1 \leq n \leq n' \tag{408}$$

Agora voltando à equação de Laguerre associada,

$$xv''(x) + \big[l_0 + 1 - x\big]v'(x) + \big[x_0 - 2(1 + l_0)\big]v(x) = 0 \tag{409}$$

Sem o fator 2 em $v'(x)$, podemos verificar o que muda no cálculo da série infinita

$$v(x) = \sum_{n=0}^{\infty} b_n x^n \tag{410}$$

na equação diferencial (409) e assim encontrar a seguinte relação de recorrência:

$$b_{n+1} = \frac{n - n'}{(n+1)(n + l_0 + 1)} b_n \tag{411}$$

com

$$x_0 = 2(1 + l_0) + n' \implies n' = x_0 - 2(1 + l_0)$$

O resultado da série (410) será um polinômio $L_{n'}^{l_0}(x)$ de grau n', e esses são chamados de polinômios associados de Laguerre.

A ausência do fator 2 na EDO (409) automaticamente nos leva aos termos gerais

$$b_n = b_0 \binom{n'}{n} \cdot \prod_{k=1}^{n} \left(-\frac{1}{l_0 + k} \right) \qquad \forall n,\ 1 \leq n \leq n'$$

ou equivalentemente

$$b_n = b_0 \frac{(-1)^n (n')!}{n!(n' - n)!} \cdot \prod_{k=1}^{n} \left(\frac{1}{l_0 + k} \right) \tag{412}$$

Podemos escolher b_0 para cada polinômio de grau n' de forma que simplifique a expressão (412)

$$b_0 = \frac{1}{(n')!} \prod_{k=1}^{n'} l_0 + k$$

de maneira que

$$b_n = \frac{(-1)^n}{n!(n' - n)!} \prod_{k=n+1}^{n'} l_0 + k$$

E o produtório também pode ser simplificado:

$$\prod_{k=n+1}^{n'} l_0 + k = (l_0 + n + 1) \cdot (l_0 + n + 2) \cdot \ldots \cdot (l_0 + n') = \frac{(l_0 + n')!}{(l_0 + n)!}$$

e então

$$b_n = \frac{(-1)^n}{n!(n' - n)!} \cdot \frac{(l_0 + n')!}{(l_0 + n)!} \tag{413}$$

Usando a relação (413) podemos encontrar os primeiros polinômios associados de Laguerre:

$$\begin{aligned}
L_0^{l_0}(x) &= 1 \\
L_1^{l_0}(x) &= -x + l_0 + 1 \\
L_2^{l_0}(x) &= \frac{1}{2}\left[x^2 - 2(l_0+2)x + (l_0+1)(l_0+2)\right] \\
L_3^{l_0}(x) &= \frac{1}{6}\left[-x^3 + 3(l_0+3)x^2 - 3(l_0+2)(l_0+3)x + (l_0+1)(l_0+2)(l_0+3)\right]
\end{aligned}$$

Comparando as relações de recorrência (406) e (411), podemos construir uma relação entre os polinômios $L_{n'}^{l_0}(x)$ e $U_{n'}^{l}(x)$. Note que se fizermos

$$n + l_0 + 1 = n + 2l + 2 \implies l_0 = 2l + 1$$

os denominadores das relações de recorrência serão idênticos. Entretanto ainda há um fator 2 na relação de $U_{n'}^{l}(x)$. Mas note que essa constante pode ser absorvida no argumento da série

$$u(x) = \sum_{n=0}^{\infty} a_n x^n \tag{414}$$

uma vez que essa constante implica na aparição do fator 2^n nos termos gerais (407). Isso significa que se colocarmos $2x$ ao invés de x no argumento do polinômio $L_{n'}^{l_0}$ (ou equivalentemente substituir x^n por $(2x)^n$ na série (410)) e fizermos $a_0 = b_0$, obteremos exatamente a mesma relação de recorrência de $U_{n'}^{l}$. Em outras palavras, temos

$$U_{n'}^{l}(x) = L_{n'}^{l_0}(2x) = L_{n'}^{2l+1}(2x) \tag{415}$$

No contexto da mecânica quântica, reservamos o caractere n para denominar o número quântico principal do átomo de hidrogênio (aqui já podemos esquecer o n como sendo o n-ésimo coeficiente do polinômio). Esse número é definido pela expressão

$$n = n' + l + 1$$

onde n' denota o grau do polinômio. Então

$$n' = n - l - 1 \tag{416}$$

Note que n' precisa ser maior ou igual a zero (já que indica o grau de um polinômio), o que implica

$$n - l - 1 \geq 0 \implies l \leq n - 1$$

ou seja, l não pode ser arbitrariamente alto.

Substituindo (416) em (415), obtemos

$$U_{n'}^{l}(x) = L_{n-l-1}^{2l+1}(2x)$$

Finalmente, a solução final para a EDO radial (em termos de x) é

$$R_{nl}(x) = x^l e^{-x} L_{n-l-1}^{2l+1}(2x) \tag{417}$$

Agora lembrando que

$$x_0 = 2(1 + l + n') = 2n$$

e que

$$\frac{2m_0 e^2}{4\pi\epsilon_0 \hbar^2} r = x_0 x$$

então

$$x = \frac{m_0 e^2}{4n\pi\epsilon_0 \hbar^2} r$$

Chamando a constante $\frac{m_0 e^2}{4\pi\epsilon_0 \hbar^2}$ de $\frac{1}{a}$, obteremos

$$x = \frac{r}{an}$$

Substituindo na equação (417), obteremos a solução radial para r:

$$R_{nl}(r) = \left(\frac{r}{an}\right)^l e^{-r/na} L_{n-l-1}^{2l+1}\left(\frac{2r}{an}\right) \tag{418}$$

A segunda solução do método de Frobenius possui um logaritmo $\ln|x|$ multiplicando um termo $u_1(x)$ que é diferente de zero em $x = 0$ (sabemos disso porque $a_0 \neq 0$, caso contrário u_1 seria a solução nula). Consequentemente essa solução terá um comportamento duvidoso nesse ponto. Manipulando a EDO (400) para a forma

$$u''(x) + 2\frac{l + 1 - x}{x} u'(x) + \frac{x_0 - 2(1 + l)}{x} u(x) = 0$$

podemos usar a identidade de Abel (6.5) no intervalo $(0, +\infty)$ para mostrar que o wronskiano $W(u_1, u_2)$ das duas soluções linearmente independentes é

$$\begin{aligned} W(u_1, u_2)(x) &= C \cdot \exp\left(-\int_{x_0}^{x} 2\frac{l + 1 - x'}{x'}\,\mathrm{d}x'\right) \\ &= C \cdot \exp\left(-2\int_{x_0}^{x} \frac{l + 1}{x'} - 1\,\mathrm{d}x'\right) \\ &= C \cdot \exp\left(-2(l + 1)\ln\left(\frac{x}{x_0}\right) + 2(x - x_0)\right) \\ &= C \cdot e^{2(x - x_0)} \left(\frac{x_0}{x}\right)^{2(l+1)} \end{aligned}$$

com $x_0 \in (0, +\infty)$ e $l \geq 0$. Então o wronskiano não converge quando $x \to 0+$, o que implica (conforme vimos na equação de Bessel esférica) que uma das duas soluções diverge em $x = 0$. Sabemos que $u_1(x)$ não diverge nesse ponto e portanto $u_2(x)$ diverge. Então descartamos essa segunda solução e ficamos apenas com a função radial (418) que encontramos anteriormente.

As primeiras soluções radiais do átomo de hidrogênio, dadas pela expressão (417) estão ilustradas nas figuras (23), (24) e (25).
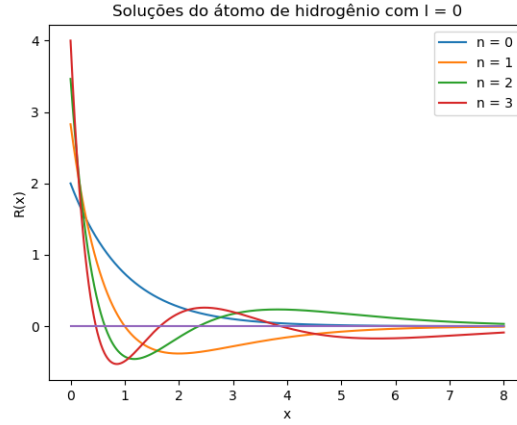
Figura 23: Primeiras quatro soluções radiais do átomo de hidrogênio com $l = 0$.
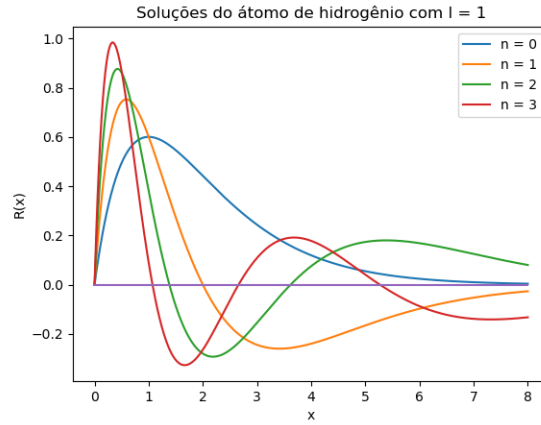
Figura 24: Primeiras quatro soluções radiais do átomo de hidrogênio com $l = 1$.

7.2 Os níveis de energia

Para encontrarmos os níveis de energia permitidos pelo átomo de hidrogênio, devemos lembrar que

$$x_0 = 2(1 + l + n') = 2n$$

E da definição de x_0:

$$x_0 = \frac{2m_0 e^2}{4\pi\epsilon_0 \hbar\sqrt{-2m_0 E}}$$
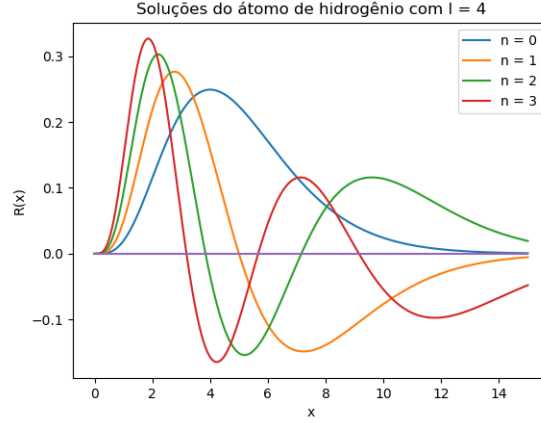
Figura 25: Primeiras quatro soluções radiais do átomo de hidrogênio com $l = 4$.

Portanto,

$$\frac{2m_0e^2}{4\pi\epsilon_0\hbar\sqrt{-2m_0E}} = 2n$$

$$-\left(\frac{2m_0e^2}{4\pi\epsilon_0\hbar}\right)^2 \cdot \frac{1}{2m_0E} = 4n^2$$

Finalmente, os níveis de energia do átomo de hidrogênio são

$$E_n = -\frac{m_0}{2\hbar^2}\left(\frac{e^2}{4\pi\epsilon_0}\right)^2 \cdot \frac{1}{n^2} \tag{419}$$

Agora note que

$$E_1 = -\frac{m_0}{2\hbar^2}\left(\frac{e^2}{4\pi\epsilon_0}\right)^2$$

o que significa que podemos reescrever a expressão da energia (419) como

$$E_n = \frac{E_1}{n^2}$$

A menor energia é a com $n = 1$, E_1. Usando os valores conhecidos da carga e massa do elétron, e e m_0 respectivamente, além da constante de Planck reduzida $\hbar$ e da permissividade elétrica no vácuo ϵ_0, obteremos

$$E_1 = -13,6\ eV$$

que é a energia necessária para ionizar um átomo de hidrogênio com seu elétron na órbita mais estável, pois o elétron se liberta do átomo com $E = 0$.

Por fim, a solução para a função de onda $\psi(r,\theta,\varphi,t)$ é dada por

$$\begin{aligned}\psi_{nlm}(r,\theta,\varphi,t) &= A_{nl}R_{nl}(r)Y_{lm}(\theta,\varphi)e^{-itE_n/\hbar}\\ &= A_{nl}\left(\frac{r}{an}\right)^l e^{-r/na}L_{n-l-1}^{2l+1}\left(\frac{2r}{an}\right)Y_{lm}(\theta,\varphi)e^{-itE_n/\hbar} \qquad (420)\end{aligned}$$

onde A_{nl} é uma constante que normaliza a função de onda ψ, n é o número quântico principal, l é denotado como número quântico orbital, m o número quântico magnético e E_n é a energia do sistema,

$$E_n = -\frac{m_0}{2\hbar^2}\left(\frac{e^2}{4\pi\epsilon_0}\right)^2 \cdot \frac{1}{n^2}$$

A solução geral é a soma das soluções particulares (420) em todos os três números quânticos (n, l e m), que corresponde à superposição de estados:

$$\psi(\vec{\mathbf{r}},t) = \sum_{n=1}^{\infty}\sum_{l=0}^{n-1}\sum_{m=-l}^{l} A_{nl}\left(\frac{r}{an}\right)^l e^{-r/na}L_{n-l-1}^{2l+1}\left(\frac{2r}{an}\right)Y_{lm}(\theta,\varphi)e^{-itE_n/\hbar}$$

Agora sabemos que

$$x_{0_n} = \frac{2m_0e^2}{4\pi\epsilon_0\hbar\sqrt{-2m_0E_n}} = \frac{2m_0e^2}{4\pi\epsilon_0\hbar\sqrt{-2m_0E_1}} \cdot n$$

Da equação diferencial

$$\frac{d}{dx}\left(x^2\frac{dR}{dx}\right) + \left[x^2 + x_{0_n}x - l(l+1)\right]R(x) = 0$$

podemos fazer $p(x) = x^2$,

$$q_1(x) = x^2 + x_{0_n}x - l(l+1) \quad \text{ e } \quad q_2(x) = x^2 + x_{0_j}x - l(l+1)$$

Se $n > j$, então $q_1 > q_2$ para qualquer $x > 0$. Então p, q_1 e q_2 satisfazem as condições do teorema da comparação de Sturm (4.3), assim como as soluções $R_n(x)$, o que implica que entre duas raízes de $R_n(x)$, haverá pelo menos uma raiz de R_{n+1}, R_{n+2}, R_{n+3} e todos os demais. Isso pode ser visto nas figuras (24) e (25).

7.3 Propriedades dos polinômios associados de Laguerre

7.3.1 Ortogonalidade e completude

A função $R_n(x) = x^l e^{-x}L_n^{2l+1}(2x)$ (aqui n volta a denotar o grau do polinômio de Laguerre associado) é solução da equação diferencial

$$x^2R''(r) + 2xR(x) + \big[-x^2 + x_0x - l(l+1)\big]R(x) = 0 \qquad (421)$$

que pode ser reescrita como

$$\frac{d}{dx}\left(x^2\frac{dR}{dx}\right)+\big[-x^2+x_0x-l(l+1)\big]R(x)=0$$

ou então

$$\frac{d}{dx}\left(p(x)\frac{dR}{dx}\right)+\big[q(x)+\lambda\omega(x)\big]R(x)=0$$

com $p(x)=x^2$, $q(x)=-(l(l+1)+x^2)$, $\lambda=x_0$ e a função peso $\omega(x)=x$. Novamente não usamos $l(l+1)$ como autovalor já que este está reservado à EDO de Legendre associada. Então $R_n(x)$ é uma autofunção de autovalor x_0 do problema de Sturm-Liouville

$$LR+\lambda\omega R=0$$

com as condições de contorno

$$R(0)\in\mathbb{R} \quad \text{e} \quad \lim_{x\to\infty}R(x)=0$$

e o produto interno

$$\langle f,g\rangle=\int_0^\infty f(x)g(x)x\,\mathrm{d}x$$

Note que n (o grau do polinômio) depende biunivocamente de x_0 pela relação

$$x_0=2(1+l+n)$$

Vemos que esse problema satisfaz a condição

$$\lim_{x\to\infty}\sqrt{\omega(x)}R_n(x)=\lim_{x\to\infty}x\cdot x^le^{-x}L_n^{2l+1}(2x)=0$$

e portanto podemos usar os resultados da teoria de Sturm-Liouville.

O teorema (4.12) nos garante que as funções $R_n(x)$ satisfazem a relação de ortogonalidade

$$\int_0^\infty R_n(x)R_m(x)x\,\mathrm{d}x=\int_0^\infty x^{2l+1}e^{-2x}L_n^{2l+1}(2x)L_m^{2l+1}(2x)\,\mathrm{d}x=c_n\delta_{nm} \quad (422)$$

onde c_n denota o quadrado da norma de $R_n(x)$.

Podemos deixar a equação (422) mais limpa se fizermos $x'=2x$, $dx'=2dx$:

$$\frac{1}{2^{2l+2}}\int_0^\infty {x'}^{2l+1}e^{-x'}L_n^{2l+1}(x')L_m^{2l+1}(x')\,\mathrm{d}x'=c_n\delta_{nm}$$

Se retirarmos a EDO (421) do contexto do problema do átomo de hidrogênio e tratá-la de forma completamente independente, podemos afirmar que o número l pode ser qualquer real, uma vez que isso não é condição para truncar as relações de recorrência (406) e (411) que geram os polinômios associados de Laguerre (a exigência l inteiro vem da EDO dos polinômios associados de Legendre). A

única mudança que teria de ser feita é substituir os fatoriais das expressões dos coeficientes a_n ou b_n que envolvem l pela função gama Γ.

Chamando $\alpha = 2l + 1$, incorporando a potência de 2 à constante c_n e mudando a variável muda x' para x, obteremos a relação de ortogonalidade dos polinômios associados de Laguerre:

$$\int_0^\infty x^\alpha e^{-x} L_n^\alpha(x) L_m^\alpha(x)\,\mathrm{d}x = d_n \delta_{nm}$$

para alguma constante d_n.

A teoria de Sturm-Liouville também nos garante que as autofunções

$$u_{nl}(x) = x^l e^{-x} L_n^{2l+1}(2x)$$

formam uma base infinita ortogonal do espaço $\mathcal{L}^2(0, \infty)$, de maneira que

$$f(x) = \sum_{n=0}^{\infty} \frac{\langle f, u_{nl}\rangle}{\langle u_{nl}, u_{nl}\rangle} u_{nl}(x) = \sum_{n=0}^{\infty} c_{nl} u_{nl}(x) \qquad \forall f \in \mathcal{L}^2(0, \infty) \tag{423}$$

com

$$c_{nl} = \frac{\langle f, u_{nl}\rangle}{\langle u_{nl}, u_{nl}\rangle} = \frac{\displaystyle\int_0^\infty f(x) x^{l+1} e^{-x} L_n^{2l+1}(2x)\,\mathrm{d}x}{\displaystyle\int_0^\infty x^{2l+1} e^{-2x} \left| L_n^{2l+1}(2x)\right|^2 \mathrm{d}x} \tag{424}$$

Exemplo 7.1 (Distribuição de Poisson no raio)**.** *Queremos que o elétron não fique perto do núcleo mas ao mesmo tempo não muito longe. Uma função plausível para a parte radial da função de onda ψ é a distribuição de Poisson,*

$$g(x) = \frac{x^k e^{-x}}{k!} \tag{425}$$

Estamos usando x ao invés de r para facilitar a compreensão, pois x só difere de r por um fator positivo. Podemos usar a expansão em autofunções dada pela fórmula (423) *para que a função $R_n(x)$ se assemelhe à função* (425)*. Usando $k = 7$, $l = 1$ e apenas seis autofunções $R_n(x)$, obteremos o gráfico da figura* (26)*. Os resultados foram obtidos usando três programas em Python: um para calcular os coeficientes da série de Taylor usando a relação de recorrência* (406) *(código no apêndice* (J.1.1)*); outro para obter as normas das autofunções do átomo de hidrogênio*(J.1.2)*; e um terceiro programa para calcular as coordenadas $\langle g, u_{nl}\rangle$ e imprimir a expansão em autofunções na tela* (J.1.3)*. Usamos apenas seis autofunções para ilustrar a convergência, pois ao usar mais do que isso, as diferenças ficam imperceptíveis.*

Exemplo 7.2 (Distribuição senoidal no raio)**.** *Agora queremos que $R_n(x)$ tenha uma aparência próxima de uma senoide,*

$$g(x) = \sin(x) \tag{426}$$
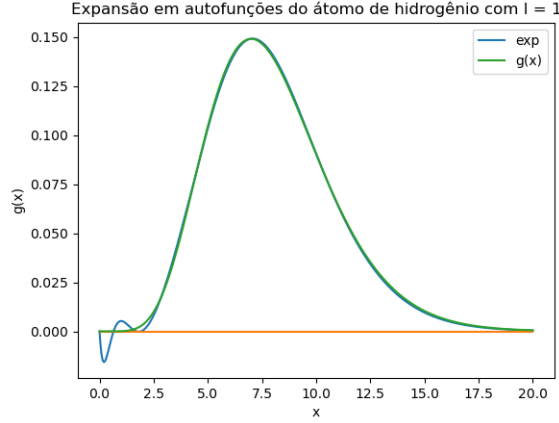
Figura 26: Expansão em autofunções da distribuição de Poisson (425) do átomo de hidrogênio usando apenas 6 autofunções com $l = 1$ e $k = 7$. A linha azul é a expansão formada pela combinação linear de seis autofunções. A linha verde é a função (425).

Apesar da função seno não ser quadrado integrável em $+\infty$, podemos usar a expressão (423) para obter a projeção da função seno no espaço gerado pelas autofunções do átomo de hidrogênio. Iremos mostrar a função não normalizada para ilustrar a convergência da expansão. Com isso, obteremos o gráfico da figura (27). Note o fenômeno de Gibbs acontecendo próximo à origem, ocorrendo grandes variações em torno de $g(x)$ até se estabilizar. As autofunções são forçadas a tenderem a zero quando x é muito grande devido ao fator exponencial e^{-x} de $R_n(x)$.

7.3.2 Relação de recorrência entre diferentes polinômios

Sejam os polinômios associados de Laguerre

$$L_n^l(x) = \sum_{k=0}^{n} b_k x^k$$

com os termos b_k calculados pela relação

$$b_k = \frac{(-1)^k}{k!(n-k)!} \cdot \frac{(l+n)!}{(l+k)!}$$

de forma que

$$L_n^l(x) = \sum_{k=0}^{n} \left[\frac{(-1)^k}{k!(n-k)!} \cdot \frac{(l+n)!}{(l+k)!} \right] x^k$$
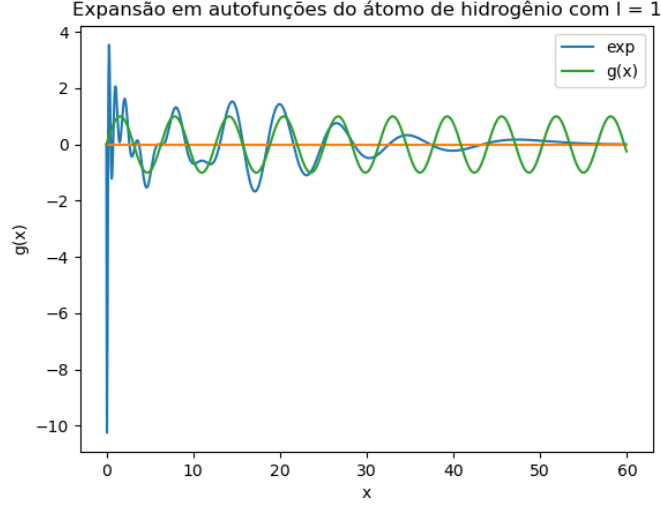
Figura 27: Expansão em autofunções do átomo de hidrogênio da função seno (426) usando vinte e cinco autofunções com $l = 1$. A linha azul é a expansão formada pela combinação linear das autofunções. A linha verde é a função seno.

Derivando essa expressão em relação a x, obteremos

$$\frac{d}{dx}L_n^l(x) = \sum_{k=1}^{n}\left[\frac{(-1)^k k}{k!(n-k)!}\cdot\frac{(l+n)!}{(l+k)!}\right]x^{k-1}$$
$$= \sum_{k=1}^{n}\left[\frac{(-1)^k}{(k-1)!(n-k)!}\cdot\frac{(l+n)!}{(l+k)!}\right]x^{k-1}$$

Fazendo a substituição de índice $k' = k-1$, temos que

$$\frac{d}{dx}L_n^l(x) = \sum_{k'=0}^{n-1}\left[\frac{(-1)^{k'+1}}{(k')!((n-1)-k')!}\cdot\frac{(l+n)!}{(l+k'+1)!}\right]x^{k'}$$

Mudando o índice mudo para k e retirando um fator -1 do somatório, temos

$$\frac{d}{dx}L_n^l(x) = -\sum_{k=0}^{n-1}\left[\frac{(-1)^{k}}{(k)!((n-1)-k)!}\cdot\frac{(l+n)!}{(l+k+1)!}\right]x^{k}$$

Mas isso é precisamente o simétrico do polinômio associado de Laguerre de grau $n-1$ e com índice $l+1$:

$$\begin{aligned}-L_{n-1}^{l+1}(x) &= -\sum_{k=0}^{n-1}\left[\frac{(-1)^k}{(k)!((n-1)-k)!}\cdot\frac{((l+1)+(n-1))!}{((l+1)+k)!}\right]x^k \\ &= -\sum_{k=0}^{n-1}\left[\frac{(-1)^k}{(k)!((n-1)-k)!}\cdot\frac{(l+n)!}{(l+k+1)!}\right]x^k\end{aligned}$$

o que implica

$$\frac{d}{dx}L_n^l(x) = -L_{n-1}^{l+1}(x) \tag{427}$$

Inserindo (427) na equação diferencial

$$x\frac{d^2}{dx^2}L_n^l(x) + \big[l+1-x\big]\frac{d}{dx}L_n^l(x) + \big[x_0 - 2(1+l)\big]L_n^l(x) = 0$$

na qual esses polinômios satisfazem, obtemos

$$xL_{n-2}^{l+2}(x) - \big[l+1-x\big]L_{n-1}^{l+1}(x) + \big[x_0 - 2(1+l)\big]L_n^l(x) = 0$$

e então

$$L_n^l(x) = \frac{\big[l+1-x\big]L_{n-1}^{l+1}(x) - xL_{n-2}^{l+2}(x)}{x_0 - 2(1+l)}$$

Mas como $x_0 = 2(1+l) + n$ implica $x_0 - 2(1+l) = n$, temos a relação de recorrência

$$L_n^l(x) = \frac{\big[l+1-x\big]L_{n-1}^{l+1}(x) - xL_{n-2}^{l+2}(x)}{n} \tag{428}$$

A equação (428) nos permite calcular os próximos polinômios sabendo apenas dois com graus consecutivos.

A Núcleo de Dirac

Teorema A.1. *Seja $D_k(x)$ a função definida pela expressão*

$$D_k(x) = \frac{1}{2L} + \frac{1}{L}\sum_{n=1}^{k}\cos\Big(\frac{n\pi x}{L}\Big) \tag{429}$$

também denotada como Núcleo de Dirac. Essa função possui as seguintes propriedades:

1. *$D_k(x)$ é uma função contínua, par e periódica de período $2L$.*
2. *$D_k(0) = \frac{1}{L}\left(k + \frac{1}{2}\right)$.*
3. *$\int_{-L}^{L} D_k(x)\,\mathrm{d}x = 1$*
4. *$D_k(x) = \frac{1}{2L}\cdot\frac{\sin\left[\left(k+\frac{1}{2}\right)\cdot\frac{\pi x}{L}\right]}{\sin\Big(\frac{\pi x}{2L}\Big)}$ se x não for um múltiplo inteiro de $2L$.*

Demonstração.

1. As funções constante e cosseno são ambas contínuas e pares, e portanto a soma delas também é contínua e par. O período T de $D_k(x)$ é determinado pelo período do primeiro cosseno da somatória, ou seja, pela diferença $\frac{\pi T}{L} - 0 = 2\pi$, e então $\frac{\pi T}{L} = 2\pi$, ou seja, $T = 2L$.

2. Fazendo $x = 0$ em (429), obtemos

$$D_k(0) = \frac{1}{2L} + \frac{1}{L}\sum_{n=1}^{k} 1 = \frac{1}{2L} + \frac{k}{L} = \frac{1}{L}\left(k + \frac{1}{2}\right)$$

3. Integrando (429):

$$\begin{aligned}
\int_{-L}^{L} D_k(x)\,\mathrm{d}x &= \int_{-L}^{L}\frac{1}{2L}\,\mathrm{d}x + \frac{1}{L}\int_{-L}^{L}\sum_{n=1}^{k}\cos\Big(\frac{n\pi x}{L}\Big)\,\mathrm{d}x \\
&= \frac{x}{2L}\Big|_{-L}^{L} + \frac{1}{L}\sum_{n=1}^{k}\int_{-L}^{L}\cos\Big(\frac{n\pi x}{L}\Big)\,\mathrm{d}x \\
&= \frac{1}{2} + \frac{1}{2} + \frac{1}{L}\sum_{n=1}^{k}\frac{L}{n\pi}\underbrace{\sin\Big(\frac{n\pi x}{L}\Big)\Big|_{-L}^{L}}_{=0} \\
&= 1
\end{aligned}$$

4. Pelas propriedades da função seno:

$$\sin(A+B) = \sin(A)\cos(B) + \cos(A)\sin(B)$$
$$\sin(A-B) = \sin(A)\cos(B) - \cos(A)\sin(B)$$

$$\sin(A+B) - \sin(A-B) = 2\cos(A)\sin(B) \tag{430}$$

Vamos escolher $A = \dfrac{n\pi x}{L}$, com $n \geq 1$, e B um valor tal que

$$A + B = C \cdot \left(n + \frac{1}{2}\right) \quad \text{e} \quad A - B = C \cdot \left(n - \frac{1}{2}\right) \tag{431}$$

para algum $C \neq 0$. Dividindo uma equação pela outra de (431), obtemos

$$\frac{A+B}{A-B} = \frac{n+\frac{1}{2}}{n-\frac{1}{2}} \tag{432}$$

Para simplificar as expressões, vamos denotar $N_n = \dfrac{n+\frac{1}{2}}{n-\frac{1}{2}}$, onde

$$N_n - 1 = \frac{1}{n-\frac{1}{2}} \quad \text{e} \quad N_n + 1 = \frac{2n}{n-\frac{1}{2}}$$

$$\frac{N_n - 1}{N_n + 1} = \frac{1}{2n}$$

Retornando à equação (432):

$$A + B = N_n \cdot (A - B) = N_n \cdot A - N_n \cdot B$$

$$B \cdot (1 + N_n) = A \cdot (N_n - 1)$$

$$B = A \cdot \frac{N_n - 1}{N_n + 1} = \frac{n\pi x}{L} \cdot \frac{1}{2n} = \frac{\pi x}{2L}$$

e então

$$A + B = \frac{n\pi x}{L} + \frac{\pi x}{2L} = \frac{\pi x}{L} \cdot \left(n + \frac{1}{2}\right)$$

$$A - B = \frac{n\pi x}{L} - \frac{\pi x}{2L} = \frac{\pi x}{L} \cdot \left(n - \frac{1}{2}\right)$$

Substituindo essas expressões em (430), obtemos

$$\sin\left(\frac{\pi x}{L} \cdot (n + \frac{1}{2})\right) - \sin\left(\frac{\pi x}{L} \cdot (n - \frac{1}{2})\right) = 2\cos\left(\frac{n\pi x}{L}\right)\sin\left(\frac{\pi x}{2L}\right)$$

Agora vamos aplicar o somatório em n em ambos os lados:

$$\sum_{n=1}^{k} \sin\left(\frac{\pi x}{L} \cdot (n + \frac{1}{2})\right) - \sin\left(\frac{\pi x}{L} \cdot (n - \frac{1}{2})\right) = 2\sin\left(\frac{\pi x}{2L}\right) \sum_{n=1}^{k} \cos\left(\frac{n\pi x}{L}\right) \tag{433}$$

Note que o primeiro membro é uma série telescópica. Se definirmos

$$a_n = \sin\left(\frac{\pi x}{L} \cdot (n - \frac{1}{2})\right)$$

então

$$\begin{aligned}\sum_{n=1}^{k} a_{n+1} - a_n &= a_{k+1} - a_1 \\ &= \sin\left(\frac{\pi x}{L} \cdot (k + 1 - \frac{1}{2})\right) - \sin\left(\frac{\pi x}{L} \cdot (1 - \frac{1}{2})\right) \\ &= \sin\left(\frac{\pi x}{L} \cdot (k + \frac{1}{2})\right) - \sin\left(\frac{\pi x}{2L}\right)\end{aligned}$$

E agora vamos isolar o somatório do cosseno em (433):

$$\sum_{n=1}^{k} \cos\left(\frac{n\pi x}{L}\right) = \frac{\sin\left(\frac{\pi x}{L} \cdot (k + \frac{1}{2})\right) - \sin\left(\frac{\pi x}{2L}\right)}{2\sin\left(\frac{\pi x}{2L}\right)}$$

$$\sum_{n=1}^{k} \cos\left(\frac{n\pi x}{L}\right) = \frac{1}{2} \cdot \frac{\sin\left(\frac{\pi x}{L} \cdot (k + \frac{1}{2})\right)}{\sin\left(\frac{\pi x}{2L}\right)} - \frac{1}{2}$$

Substituindo esse somatório na definição do núcleo de Dirichlet (429), obtemos

$$D_k(x) = \frac{1}{2L} + \frac{1}{L} \cdot \left[\frac{1}{2} \cdot \frac{\sin\left(\frac{\pi x}{L} \cdot (k + \frac{1}{2})\right)}{\sin\left(\frac{\pi x}{2L}\right)} - \frac{1}{2}\right]$$

$$D_k(x) = \frac{1}{2L} \cdot \frac{\sin\left[\left(k + \frac{1}{2}\right) \cdot \frac{\pi x}{L}\right]}{\sin\left(\frac{\pi x}{2L}\right)}$$

como queríamos demonstrar.

■

B Lema de Riemann-Lebesgue

Antes de demonstrarmos o lema de Riemann-Lebesgue, precisaremos provar um outro teorema:

Teorema B.1. *Se f é uma função real que é módulo integrável em um intervalo $[a, b]$, então $\forall \epsilon > 0$ existe uma função contínua $\psi(x)$, também definida em $[a, b]$, tal que*

$$\int_a^b |f(x) - \psi(x)| \, \mathrm{d}x < \epsilon$$

e ao mesmo tempo

$$\psi(a) = \psi(b) = 0$$

Demonstração. Suponha que f seja limitada e integrável em $[a, b]$. Isso significa que existe uma função escada inferior χ definida no mesmo intervalo pela expressão

$$\chi(x) = m_i \quad \text{se} \quad x_{i-1} \leq x < x_i \quad \text{com} \quad x_0 = a \quad \text{e} \quad x_k = b$$

$$m_i \text{ é o ínfimo de } f \text{ no intervalo } [x_{i-1}, x_i]$$

tal que sua integral $\displaystyle\int_a^b \chi(x) \, \mathrm{d}x = \sum_{i=1}^{k} m_i(x_i - x_{i-1})$ satisfaz

$$\int_a^b f(x) \, \mathrm{d}x - \int_a^b \chi(x) \, \mathrm{d}x < \frac{\epsilon}{2}$$

ou equivalentemente

$$\int_a^b (f(x) - \chi(x)) \, \mathrm{d}x < \frac{\epsilon}{2} \tag{434}$$

Agora seja $\psi_y(x)$ uma função definida em $[a, b]$, onde $y \in [0, \frac{\pi}{2}]$, tal que

$$\text{se } x \in [x_{i-1}, x_i]$$

$$\psi_y(x) = \begin{cases} \tan(y_i) \cdot [x - x_{i-1}] & \text{se } x \in [x_{i-1}, v_{i-1}] \\ m_i & \text{se } x \in [v_{i-1}, w_i] \\ \tan(y_i) \cdot [x_i - x] & \text{se } x \in [w_i, x_i] \end{cases}$$

onde

$$y_i = \begin{cases} y & \text{se } m_i \geq 0 \\ -y & \text{se } m_i < 0 \end{cases}$$

$$v_{i-1} = x_{i-1} + |m_i| \cdot \cot(y_i) \quad \text{e} \quad w_i = x_i - |m_i| \cdot \cot(y_i)$$

Em outras palavras, $\psi_y(x)$ é uma função que gera trapézios em cada subintervalo $[x_{i-1}, x_i]$ e os lados inclinados desses trapézios fazem um ângulo de y radianos em relação ao eixo x. Portanto a função

$$|\chi(x) - \psi_y(x)|$$

quando não nula forma triângulos retângulos cuja área no subintervalo $[x_{i-1}, x_i]$ é

$$\frac{|m_i| \cdot (v_{i-1} - x_i)}{2} + \frac{|m_i| \cdot (x_i - w_i)}{2} = |m_i|^2 \cot(y)$$

uma vez que $|y_i| = y$. E então podemos afirmar que

$$\int_a^b |\chi(x) - \psi_y(x)| \, \mathrm{d}x = \sum_{i=1}^{k} |m_i|^2 \cot(y)$$

Uma vez que f é uma função limitada em $[a, b]$, existe $M > 0$ tal que $|f(x)| \leq M$. Isso significa que $|m_i| \leq M \quad \forall i$ e então

$$\int_a^b |\chi(x) - \psi_y(x)| \, \mathrm{d}x \leq \sum_{i=1}^{k} M^2 \cot(y) = k \cdot M^2 \cot(y)$$

Uma vez que k é finito e a função cotangente é contínua no intervalo $(0, \frac{\pi}{2}]$, para todo $\epsilon > 0$ sempre podemos encontrar um y próximo de $\frac{\pi}{2}$ tal que

$$\cot(y) < \frac{\epsilon}{2kM^2}$$

e portanto

$$\int_a^b |\chi(x) - \psi_y(x)| \, \mathrm{d}x < \frac{\epsilon}{2}$$

Lembrando que como $\chi(x)$ é uma função escada inferior de f, então $|f(x) - \chi(x)| = f(x) - \chi(x)$ em $[a, b]$ e então podemos concluir que

$$\begin{aligned}
\int_a^b |f(x) - \psi_y(x)| \, \mathrm{d}x &= \int_a^b |f(x) - \chi(x) + \chi(x) - \psi_y(x)| \, \mathrm{d}x \\
&\leq \int_a^b |f(x) - \chi(x)| + |\chi(x) - \psi_y(x)| \, \mathrm{d}x \\
&\leq \int_a^b |f(x) - \chi(x)| \, \mathrm{d}x + \int_a^b |\chi(x) - \psi_y(x)| \, \mathrm{d}x \\
&< \frac{\epsilon}{2} + \frac{\epsilon}{2} \\
&< \epsilon
\end{aligned}$$

e ainda

$$\begin{aligned}
\psi_y(a) = \tan(y_i) \cdot [x_0 - x_0] = 0 \\
\psi_y(b) = \tan(y_i) \cdot [x_k - x_k] = 0
\end{aligned}$$

o que prova o teorema no caso particular em que f é limitada.

Agora suponha que f não seja limitada apenas nas extremidades do intervalo $[a, b]$ mas ainda seja integrável e integrável em módulo nesse intervalo como uma integral imprópria. Então dado $\epsilon > 0$ existe algum $\delta > 0$ tal que

$$\left| \int_a^b |f(x)| \, \mathrm{d}x - \int_{a+\delta}^{b-\delta} |f(x)| \, \mathrm{d}x \right| < \frac{\epsilon}{2} \tag{435}$$

dado que f é limitada em $[a+\delta, b-\delta]$. Consequentemente, podemos usar o teorema (B.1) em seu caso particular nesse intervalo e portanto existe alguma função contínua $\psi(x)$ definida em $[a+\delta, b-\delta]$ tal que

$$\int_{a+\delta}^{b-\delta} |f(x) - \psi(x)| \, \mathrm{d}x < \frac{\epsilon}{2} \tag{436}$$

e $\psi(a+\delta) = \psi(b-\delta) = 0$.

Agora vamos expandir a função $\psi(x)$ da seguinte forma: vamos definir uma outra função ψ_+ no intervalo $[a, b]$

$$\psi_+(x) = \begin{cases} \psi(x) & \text{se } x \in [a+\delta, b-\delta] \\ 0 & \text{se } x \in [a, a+\delta] \text{ ou } x \in [b-\delta, b] \end{cases}$$

E então

$$\int_a^b |f(x) - \psi_+(x)| \, \mathrm{d}x = \int_a^{a+\delta} |f(x) - 0| \, \mathrm{d}x + \int_{a+\delta}^{b-\delta} |f(x) - \psi(x)| \, \mathrm{d}x + \int_{b-\delta}^{b} |f(x) - 0| \, \mathrm{d}x \tag{437}$$

Agora note que

$$\left| \int_a^b |f(x)| \, \mathrm{d}x - \int_{a+\delta}^{b-\delta} |f(x)| \, \mathrm{d}x \right| = \left| \int_a^{a+\delta} |f(x)| \, \mathrm{d}x + \int_{b-\delta}^{b} |f(x)| \, \mathrm{d}x \right|$$

Como ambas as integrais são não negativas, podemos retirar o módulo e afirmar que

$$\left| \int_a^b |f(x)| \, \mathrm{d}x - \int_{a+\delta}^{b-\delta} |f(x)| \, \mathrm{d}x \right| = \int_a^{a+\delta} |f(x)| \, \mathrm{d}x + \int_{b-\delta}^{b} |f(x)| \, \mathrm{d}x < \frac{\epsilon}{2}$$

E usando essa desigualdade e a de (436) em (437), concluímos que

$$\int_a^b |f(x) - \psi_+(x)| \, \mathrm{d}x < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

com $\psi_+(a) = \psi_+(b) = 0$, o que completa a demonstração para qualquer f integrável e absolutamente integrável. ■

Note que se f tiver outros pontos em que é ilimitada em $[a, b]$ que não sejam as extremidades, podemos repartir em subintervalos em que f só seja ilimitada nas extremidades de cada subintervalo e aplicar o teorema.

E agora podemos provar o Lema de Riemann-Lebesgue.

Teorema B.2 (Lema de Riemann-Lebesgue)**.** *Se f é uma função integrável e integrável em módulo em um intervalo $[a, b]$, então*

$$\lim_{k\to\infty} \int_a^b f(x) \cdot \sin(kx) \, \mathrm{d}x = 0 \qquad e \qquad \lim_{k\to\infty} \int_a^b f(x) \cdot \cos(kx) \, \mathrm{d}x = 0$$

Demonstração. Suponha primeiramente que f seja limitada em $[a, b]$. Então existem uma função escada superior τ e uma função escada inferior χ definidas como

$$\tau(x) = M_i \quad \text{se} \quad x_{i-1} \leq x < x_i \quad \text{com} \quad x_0 = a \quad \text{e} \quad x_n = b$$

$$M_i \text{ é o supremo de } f \text{ no intervalo } [x_{i-1}, x_i]$$

$$\chi(x) = m_i \quad \text{se} \quad x_{i-1} \leq x < x_i \quad \text{com} \quad x_0 = a \quad \text{e} \quad x_n = b$$

$$m_i \text{ é o ínfimo de } f \text{ no intervalo } [x_{i-1}, x_i]$$

que satisfazem

$$\int_a^b \tau(x)\,\mathrm{d}x - \int_a^b \chi(x)\,\mathrm{d}x = \int_a^b (\tau(x) - \chi(x))\,\mathrm{d}x = \sum_{i=1}^{n} (M_i - m_i) \cdot (x_i - x_{i-1}) < \frac{\epsilon}{2} \tag{438}$$

para um $\epsilon > 0$ arbitrário. Além disso, existe $M > 0$ tal que $|f(x)| \leq M$ em todo o intervalo $[a, b]$.

A demonstração do lema de Riemann-Lebesgue para o cosseno é análogo ao do seno, pois não faz nenhuma diferença substancial nas contas que serão feitas a seguir. Usando a mesma partição usada na definição de τ e χ, podemos reescrever a integral $\int_a^b f(x) \sin(kx)\,\mathrm{d}x$ como

$$\int_a^b f(x) \sin(kx)\,\mathrm{d}x = \sum_{i=1}^{n} \left[\int_{x_{i-1}}^{x_i} f(x) \sin(kx)\,\mathrm{d}x \right]$$

Somando $f(x_i) \cdot \int_{x_{i-1}}^{x_i} \sin(kx)\,\mathrm{d}x - f(x_i) \cdot \int_{x_{i-1}}^{x_i} \sin(kx)\,\mathrm{d}x = 0$ em cada termo do somatório resulta em

$$\int_a^b f(x) \sin(kx)\,\mathrm{d}x = \sum_{i=1}^{n} \left[\int_{x_{i-1}}^{x_i} f(x) \sin(kx)\,\mathrm{d}x \right] + f(x_i) \cdot \int_{x_{i-1}}^{x_i} \sin(kx)\,\mathrm{d}x - f(x_i) \cdot \int_{x_{i-1}}^{x_i} \sin(kx)\,\mathrm{d}x$$

$$\int_a^b f(x) \sin(kx)\,\mathrm{d}x = \sum_{i=1}^{n} \left[f(x_i) \cdot \int_{x_{i-1}}^{x_i} \sin(kx)\,\mathrm{d}x \right] + \sum_{i=1}^{n} \left[\int_{x_{i-1}}^{x_i} (f(x) - f(x_i)) \cdot \sin(kx)\,\mathrm{d}x \right]$$

Note que

$$\left|\int_{x_{i-1}}^{x_i} \sin(kx)\,\mathrm{d}x\right| = \frac{1}{k} \cdot \left| -\cos(kx)\Big|_{x_{i-1}}^{x_i} \right|$$
$$= \frac{1}{k} \cdot |\cos(kx_i) - \cos(kx_{i-1})|$$
$$\leq \frac{2}{k}$$

além disso, da desigualdade $|f(x)| \leq M$ em todo o intervalo $[a, b]$, é trivial que

$$|f(x_i)| \leq M$$

a função seno é limitada:

$$|\sin(kx)| \leq 1$$

e para todo $x \in [x_{i-1}, x_i]$, temos

$$|f(x) - f(x_i)| \leq M_i - m_i$$

visto que $M_i - m_i$ é a diferença máxima entre os valores de f nesse intervalo. Com isso, podemos afirmar que

$$\left|\int_a^b f(x)\sin(kx)\,\mathrm{d}x\right| \leq \sum_{i=1}^{n} |f(x_i)| \cdot \left|\int_{x_{i-1}}^{x_i} \sin(kx)\,\mathrm{d}x\right| + \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} |f(x) - f(x_i)| \cdot |\sin(kx)|\,\mathrm{d}x$$

$$\left|\int_a^b f(x)\sin(kx)\,\mathrm{d}x\right| \leq \sum_{i=1}^{n} M \cdot \frac{2}{k} + \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} (M_i - m_i)\,\mathrm{d}x$$
$$\leq \frac{2nM}{k} + \sum_{i=1}^{n} (M_i - m_i) \cdot (x_i - x_{i-1})$$

Agora note que esse somatório é o mesmo que aparece em (438):

$$\sum_{i=1}^{n} (M_i - m_i) \cdot (x_i - x_{i-1}) < \frac{\epsilon}{2}$$

Além disso, para todo $\epsilon > 0$, podemos encontrar um K tal que

$$\frac{2nM}{k} < \frac{\epsilon}{2}$$

para todo $k \geq K$. Isso significa que

$$\left|\int_a^b f(x)\sin(kx)\,\mathrm{d}x\right| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

Em outras palavras,

$$\lim_{k\to\infty}\int_a^b f(x)\cdot\sin(kx)\,\mathrm{d}x = 0$$

o que prova o teorema no caso particular em que f é limitada em $[a,b]$.

Agora suponha que f não necessariamente seja limitada em $[a,b]$, mas ainda seja integrável e absolutamente integrável nesse intervalo. O teorema (B.1) garante que existe alguma função contínua ψ definida em $[a,b]$ tal que

$$\int_a^b |f(x)-\psi(x)|\,\mathrm{d}x < \frac{\epsilon}{2}$$

Como a função $\psi(x)$ é contínua em $[a,b]$, então ela é limitada nesse intervalo e podemos usar o lema de Riemann-Lebesgue para $\psi(x)$ e afirmar que

$$\left|\int_a^b \psi(x)\sin(kx)\,\mathrm{d}x\right| < \frac{\epsilon}{2}$$

Agora seja a integral $\int_a^b f(x)\sin(kx)\,\mathrm{d}x$. Somando o termo

$$\int_a^b \psi(x)\sin(kx)\,\mathrm{d}x - \int_a^b \psi(x)\sin(kx)\,\mathrm{d}x = 0$$

à integral, obtemos

$$\int_a^b f(x)\sin(kx)\,\mathrm{d}x = \int_a^b f(x)\sin(kx)\,\mathrm{d}x + \\ \int_a^b \psi(x)\sin(kx)\,\mathrm{d}x - \int_a^b \psi(x)\sin(kx)\,\mathrm{d}x$$

$$\int_a^b f(x)\sin(kx)\,\mathrm{d}x = \int_a^b \psi(x)\sin(kx)\,\mathrm{d}x + \int_a^b [f(x)-\psi(x)]\sin(kx)\,\mathrm{d}x$$

Tomando o valor absoluto, concluímos que

$$\begin{aligned}\left|\int_a^b f(x)\sin(kx)\,\mathrm{d}x\right| &\leq \left|\int_a^b \psi(x)\sin(kx)\,\mathrm{d}x\right| + \int_a^b |f(x)-\psi(x)|\cdot|\sin(kx)|\,\mathrm{d}x \\ &< \frac{\epsilon}{2} + \int_a^b |f(x)-\psi(x)|\,\mathrm{d}x \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} \\ &< \epsilon\end{aligned}$$

o que completa a demonstração do lema de Riemann-Lebesgue para qualquer função integrável e integrável em módulo. ■

C Unicidade de soluções de equações diferenciais de segunda ordem

Teorema C.1. *Seja a equação diferencial do tipo*

$$f''(x) + a_1(x)f'(x) + a_0(x)f(x) = g(x) \tag{439}$$

com as condições de contorno

$$\begin{aligned} \alpha_1 f(a) + \beta_1 f'(a) = c_1 \quad (\alpha_1; \beta_1) \neq (0; 0) \\ \alpha_2 f(b) + \beta_2 f'(b) = c_2 \quad (\alpha_2; \beta_2) \neq (0; 0) \end{aligned}$$

onde $\alpha_1, \alpha_2, \beta_1, \beta_2, c_1$ e c_2 são coeficientes reais e a_0, a_1 e g são funções reais contínuas em $[a, b]$. A solução existe e é única se e somente se existirem duas funções linearmente independentes u_1 e u_2 tais que são soluções da equação homogênea

$$\begin{aligned} u_1''(x) + a_1(x)u_1'(x) + a_0(x)u_1(x) = 0 \\ u_2''(x) + a_1(x)u_2'(x) + a_0(x)u_2(x) = 0 \end{aligned}$$

obedecendo as condições de contorno e simultaneamente

$$\det \begin{bmatrix} \alpha_1 u_1(a) + \beta_1 u_1'(a) & \alpha_1 u_2(a) + \beta_1 u_2'(a) \\ \alpha_1 u_1(b) + \beta_1 u_1'(b) & \alpha_1 u_2(b) + \beta_1 u_2'(b) \end{bmatrix} \neq 0$$

Demonstração. Sabemos que a solução geral de (439) pode ser escrita como

$$f(x) = A_1 u_1(x) + A_2 u_2(x) + f_p(x) \tag{440}$$

onde u_1 e u_2 são soluções da equação homogênea $u''(x)+a_1(x)u'(x)+a_0(x)u(x) = 0$, A_1 e A_2 são duas constantes arbitrárias e $f_p(x)$ é uma solução particular de (439). Substituindo f nas condições de contorno, obtemos

$$\begin{aligned} \alpha_1[A_1 u_1(a) + A_2 u_2(a) + f_p(a)] + \beta_1[A_1 u_1'(a) + A_2 u_2'(a) + f_p'(a)] = c_1 \\ \alpha_2[A_1 u_1(b) + A_2 u_2(b) + f_p(b)] + \beta_2[A_1 u_1'(b) + A_2 u_2'(b) + f_p'(b)] = c_2 \end{aligned}$$

Colocando os coeficientes A_1 e A_2 em evidência e colocando os termos com as soluções particulares para o segundo membro, obtemos

$$\begin{aligned} A_1[\alpha_1 u_1(a) + \beta_1 u_1'(a)] + A_2[\alpha_1 u_2(a) + \beta_1 u_2'(a)] = c_1 - \alpha_1 f_p(a) - \beta_1 f_p'(a) \\ A_1[\alpha_2 u_1(b) + \beta_2 u_1'(b)] + A_2[\alpha_2 u_2(b) + \beta_2 u_2'(b)] = c_2 - \alpha_2 f_p(b) - \beta_2 f_p'(b) \end{aligned}$$

No formato matricial, temos

$$\begin{bmatrix} \alpha_1 u_1(a) + \beta_1 u_1'(a) & \alpha_1 u_2(a) + \beta_1 u_2'(a) \\ \alpha_2 u_1(b) + \beta_2 u_1'(b) & \alpha_2 u_2(b) + \beta_2 u_2'(b) \end{bmatrix} . \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = \begin{bmatrix} c_1 - \alpha_1 f_p(a) - \beta_1 f_p'(a) \\ c_2 - \alpha_2 f_p(b) - \beta_2 f_p'(b) \end{bmatrix}$$

O teorema (1.37) garante que esse sistema de equações matricial só possui uma única solução se e somente se

$$\det\begin{bmatrix} \alpha_1 u_1(a) + \beta_1 u_1'(a) & \alpha_1 u_2(a) + \beta_1 u_2'(a) \\ \alpha_1 u_1(b) + \beta_1 u_1'(b) & \alpha_1 u_2(b) + \beta_1 u_2'(b) \end{bmatrix} \neq 0$$

Se esse for o caso, A_1 e A_2 são únicos e então a solução (440) é única. ∎

Teorema C.2. *Seja a equação diferencial* (439) *e também as constantes reais* $\alpha_1, \alpha_2, \beta_1$ *e* β_2 *tais que*

$$(\alpha_1; \beta_1) \neq (0; 0) \quad e \quad (\alpha_2; \beta_2) \neq (0; 0)$$

Se v_1 *e* v_2 *são duas soluções linearmente independentes* (439) *definidas no intervalo* $[a, b]$ *e, além disso,* u_1 *e* u_2 *também são duas soluções linearmente independentes da mesma equação no mesmo intervalo* $[a, b]$*, então*

$$\det\begin{bmatrix} \alpha_1 v_1(a) + \beta_1 v_1'(a) & \alpha_1 v_2(a) + \beta_1 v_2'(a) \\ \alpha_2 v_1(b) + \beta_2 v_1'(b) & \alpha_2 v_2(b) + \beta_2 v_2'(b) \end{bmatrix} \neq 0 \tag{441}$$

se e somente se

$$\det\begin{bmatrix} \alpha_1 u_1(a) + \beta_1 u_1'(a) & \alpha_1 u_2(a) + \beta_1 u_2'(a) \\ \alpha_2 u_1(b) + \beta_2 u_1'(b) & \alpha_2 u_2(b) + \beta_2 u_2'(b) \end{bmatrix} \neq 0 \tag{442}$$

Demonstração. Ambos os conjuntos $\{v_1, v_2\}$ e $\{u_1, u_2\}$ formam uma base para o espaço das soluções da equação (439), uma vez que a dimensão desse espaço para EDOs de segunda ordem é 2. Isso significa que podemos escrever uma solução em termos de outras duas:

$$\begin{aligned} v_1(x) &= c_{11} u_1(x) + c_{12} u_2(x) \\ v_2(x) &= c_{21} u_1(x) + c_{22} u_2(x) \end{aligned}$$

No formato matricial, temos

$$\begin{bmatrix} v_1(x) \\ v_2(x) \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \cdot \begin{bmatrix} u_1(x) \\ u_2(x) \end{bmatrix}$$

Note que as linhas da matriz quadrada precisam ser linearmente independentes entre si, pois caso contrário poderíamos multiplicar uma linha por uma constante e verificar que v_1 e v_1 seriam linearmente dependentes, o que contradiz à hipótese. Portanto podemos usar o teorema (1.34) que garante

$$\det\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \neq 0 \tag{443}$$

Multiplicando essa matriz quadrada de coeficientes pela transposta de (442), obtemos

$$\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \cdot \begin{bmatrix} \alpha_1 u_1(a) + \beta_1 u_1'(a) & \alpha_2 u_1(b) + \beta_2 u_1'(b) \\ \alpha_1 u_2(a) + \beta_1 u_2'(a) & \alpha_2 u_2(b) + \beta_2 u_2'(b) \end{bmatrix} = \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{bmatrix}$$

onde

$$\begin{aligned} d_{11} &= \alpha_1 c_{11} u_1(a) + \beta_1 c_{11} u_1'(a) + \alpha_1 c_{12} u_2(a) + \beta_1 c_{12} u_2'(a) \\ &= \alpha_1 [c_{11} u_1(a) + c_{12} u_2(a)] + \beta_1 [c_{11} u_1'(a) + c_{12} u_2'(a)] \\ &= \alpha_1 v_1(a) + \beta_1 v_1'(a) \end{aligned}$$

$$\begin{aligned} d_{12} &= \alpha_2 c_{11} u_1(b) + \beta_2 c_{11} u_1'(b) + \alpha_2 c_{12} u_2(b) + \beta_2 c_{12} u_2'(b) \\ &= \alpha_2 [c_{11} u_1(b) + c_{12} u_2(b)] + \beta_2 [c_{11} u_1'(b) + c_{12} u_2'(b)] \\ &= \alpha_2 v_1(b) + \beta_2 v_1'(b) \end{aligned}$$

$$\begin{aligned} d_{21} &= \alpha_1 c_{21} u_1(a) + \beta_1 c_{21} u_1'(a) + \alpha_1 c_{22} u_2(a) + \beta_1 c_{22} u_2'(a) \\ &= \alpha_1 [c_{21} u_1(a) + c_{22} u_2(a)] + \beta_1 [c_{21} u_1'(a) + c_{22} u_2'(a)] \\ &= \alpha_1 v_2(a) + \beta_1 v_2'(a) \end{aligned}$$

$$\begin{aligned} d_{22} &= \alpha_2 c_{21} u_1(b) + \beta_2 c_{21} u_1'(b) + \alpha_2 c_{22} u_2(b) + \beta_2 c_{22} u_2'(b) \\ &= \alpha_2 [c_{21} u_1(b) + c_{22} u_2(b)] + \beta_2 [c_{21} u_1'(b) + c_{22} u_2'(b)] \\ &= \alpha_2 v_2(b) + \beta_2 v_2'(b) \end{aligned}$$

Com isso, perceba que a matriz resultante $[d_{ij}]$ é a transposta de (441). Uma vez que o determinante da matriz de coeficientes (443) é diferente de zero e que o determinante da transposta de uma matriz é o determinante da própria matriz, segue que o determinante de (441) só pode ser diferente de zero se e somente se o determinante de (442) também for diferente de zero. ■

D Soluções da equação homogênea do operador de Sturm-Liouville

Teorema D.1. *Se existirem duas soluções u_1 e u_2 da equação homogênea que obedecem à condição de unicidade* (246)*, então também existem duas funções v_1 e v_2 linearmente independentes definidas em $[a, b]$ que satisfazem $Lv_1 = Lv_2 = 0$,*

$$\alpha_1 v_1(a) + \beta_1 v_1'(a) = 0 \tag{444}$$

$$\alpha_2 v_2(b) + \beta_2 v_2'(b) = 0 \tag{445}$$

e ainda

$$W(v_1, v_2) = \det \begin{bmatrix} v_1(x) & v_2(x) \\ v_1'(x) & v_2'(x) \end{bmatrix} \neq 0 \quad \forall x \in [a, b]$$

$$\begin{aligned} \alpha_1 v_2(a) + \beta_1 v_2'(a) &\neq 0 \\ \alpha_2 v_1(b) + \beta_2 v_1'(b) &\neq 0 \end{aligned}$$

Demonstração. Sejam u_1 e u_2 duas soluções da equação homogênea $Lu = 0$ que satisfazem a condição de unicidade. Vamos definir uma matriz quadrada $[c_{ij}]$ multiplicando (246) por uma outra matriz:

$$\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} = \begin{bmatrix} \alpha_1 u_1(a) + \beta_1 u_1'(a) & \alpha_1 u_2(a) + \beta_1 u_2'(a) \\ \alpha_2 u_1(b) + \beta_2 u_1'(b) & \alpha_2 u_2(b) + \beta_2 u_2'(b) \end{bmatrix} \cdot \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} \alpha_1 u_2(a) + \beta_1 u_2'(a) & -\alpha_1 u_1(a) - \beta_1 u_1'(a) \\ \alpha_2 u_2(b) + \beta_2 u_2'(b) & -\alpha_2 u_1(b) - \beta_2 u_1'(b) \end{bmatrix} \tag{446}$$

Uma vez que ambas as matrizes do produto tem determinante diferente de zero, segue que o determinante de $\{c_{ij}\}$ também é diferente de zero:

$$\det \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \neq 0 \tag{447}$$

Agora sejam as funções v_1 e v_2 dadas por

$$\begin{bmatrix} v_1(x) & 0 \\ v_2(x) & 0 \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \cdot \begin{bmatrix} u_1(x) & 0 \\ u_2(x) & 0 \end{bmatrix} \tag{448}$$

que é equivalente ao par de equações

$$v_1(x) = c_{11} u_1(x) + c_{12} u_2(x) \tag{449}$$

$$v_2(x) = c_{21} u_1(x) + c_{22} u_2(x) \tag{450}$$

Aplicando o operador de Sturm-Liouville a ambas equações, obtemos a primeira propriedade de v_1 e v_2:

$$Lv_1(x) = c_{11} Lu_1(x) + c_{12} Lu_2(x) = 0$$
$$Lv_2(x) = c_{21} Lu_1(x) + c_{22} Lu_2(x) = 0$$

Do par de equações (449) e (450), também temos que

$$\begin{bmatrix} 0 & v_1'(x) \\ 0 & v_2'(x) \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \cdot \begin{bmatrix} 0 & u_1'(x) \\ 0 & u_2'(x) \end{bmatrix} \tag{451}$$

Somando as equações matriciais (448) e (451), obtemos

$$\begin{bmatrix} v_1(x) & v_1'(x) \\ v_2(x) & v_2'(x) \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \cdot \begin{bmatrix} u_1(x) & u_1'(x) \\ u_2(x) & u_2'(x) \end{bmatrix} \tag{452}$$

Perceba que a matriz do primeiro membro de (452) é a transposta da wronskiana das funções v_1 e v_2, assim como a segunda matriz do segundo membro é a transposta da wronskiana das funções u_1 e u_2. Como u_1 e u_2 são linearmente independentes por hipótese, segue que o determinante da wronskiana das funções u_1 e u_2 é diferente de zero. Além disso, o teorema (1.42) garante que essa desigualdade seja verdadeira para todo x em $[a, b]$. Com isso, sabendo que o determinante de uma matriz não muda com sua transposta, e que o determinante

dos coeficientes é diferente de zero (447), obtemos a terceira propriedade de v_1 e v_2 (assim como a independência linear entre as duas funções):

$$W(v_1, v_2) = \det \begin{bmatrix} v_1(x) & v_2(x) \\ v_1'(x) & v_2'(x) \end{bmatrix} \neq 0 \quad \forall x \in [a, b] \tag{453}$$

Agora vamos multiplicar a matriz do primeiro membro de (452) pela matriz coluna dos coeficientes α_1 e β_1:

$$\begin{bmatrix} v_1(x) & v_1'(x) \\ v_2(x) & v_2'(x) \end{bmatrix} \cdot \begin{bmatrix} \alpha_1 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} \alpha_1 v_1(x) + \beta_1 v_1'(x) \\ \alpha_1 v_2(x) + \beta_1 v_2'(x) \end{bmatrix} \tag{454}$$

Mas também, pela equação (452), temos

$$\begin{aligned} \begin{bmatrix} v_1(x) & v_1'(x) \\ v_2(x) & v_2'(x) \end{bmatrix} \cdot \begin{bmatrix} \alpha_1 \\ \beta_1 \end{bmatrix} &= \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \cdot \begin{bmatrix} u_1(x) & u_1'(x) \\ u_2(x) & u_2'(x) \end{bmatrix} \cdot \begin{bmatrix} \alpha_1 \\ \beta_1 \end{bmatrix} \\ &= \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \cdot \begin{bmatrix} \alpha_1 u_1(x) + \beta_1 u_1'(x) \\ \alpha_1 u_2(x) + \beta_1 u_2'(x) \end{bmatrix} \end{aligned}$$

Portanto, fazendo $x = a$, obtemos

$$\begin{bmatrix} \alpha_1 v_1(a) + \beta_1 v_1'(a) \\ \alpha_1 v_2(a) + \beta_1 v_2'(a) \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \cdot \begin{bmatrix} \alpha_1 u_1(a) + \beta_1 u_1'(a) \\ \alpha_1 u_2(a) + \beta_1 u_2'(a) \end{bmatrix}$$

Comparando os valores da matriz coluna do segundo membro com os coeficientes de (446), concluímos que

$$\begin{aligned} \begin{bmatrix} \alpha_1 v_1(a) + \beta_1 v_1'(a) \\ \alpha_1 v_2(a) + \beta_1 v_2'(a) \end{bmatrix} &= \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \cdot \begin{bmatrix} -c_{12} \\ c_{11} \end{bmatrix} \\ &= \begin{bmatrix} -c_{11}c_{12} + c_{12}c_{11} \\ -c_{21}c_{12} + c_{22}c_{11} \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ c_{22}c_{11} - c_{21}c_{12} \end{bmatrix} \end{aligned} \tag{455}$$

A primeira linha desta última equação matricial nos diz, então, que

$$\alpha_1 v_1(a) + \beta_1 v_1'(a) = 0$$

Agora iremos repetir os mesmos passos da equação (454) mas usando a matriz coluna dos coeficientes α_2 e β_2:

$$\begin{bmatrix} v_1(x) & v_1'(x) \\ v_2(x) & v_2'(x) \end{bmatrix} \cdot \begin{bmatrix} \alpha_2 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} \alpha_2 v_1(x) + \beta_2 v_1'(x) \\ \alpha_2 v_2(x) + \beta_2 v_2'(x) \end{bmatrix} \tag{456}$$

Isso significa que podemos encontrar as equações seguintes apenas trocando α_1 por α_2 e β_1 por β_2. Neste caso, vamos usar $x = b$ ao invés de $x = a$:

$$\begin{bmatrix} \alpha_2 v_1(b) + \beta_2 v_1'(b) \\ \alpha_2 v_2(b) + \beta_2 v_2'(b) \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \cdot \begin{bmatrix} \alpha_2 u_1(b) + \beta_2 u_1'(b) \\ \alpha_2 u_2(b) + \beta_2 u_2'(b) \end{bmatrix}$$

Novamente comparando os valores da matriz coluna do segundo membro com os coeficientes de (446), temos

$$\begin{aligned}
\begin{bmatrix} \alpha_2 v_1(b) + \beta_2 v_1'(b) \\ \alpha_2 v_2(b) + \beta_2 v_2'(b) \end{bmatrix} &= \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \cdot \begin{bmatrix} -c_{22} \\ c_{21} \end{bmatrix} \\
&= \begin{bmatrix} -c_{11}c_{22} + c_{12}c_{21} \\ -c_{21}c_{22} + c_{22}c_{21} \end{bmatrix} \\
&= \begin{bmatrix} c_{12}c_{21} - c_{11}c_{22} \\ 0 \end{bmatrix} \qquad (457)
\end{aligned}$$

o que significa

$$\alpha_2 v_2(b) + \beta_2 v_2'(b) = 0$$

Portanto a segunda propriedade das funções v_1 e v_2 descrita pelas equações (444) e (445) está demonstrada.

Além disso, note que os termos resultantes das equações (455) e (457)

$$\begin{aligned}
c_{22}c_{11} - c_{21}c_{12} \\
c_{12}c_{21} - c_{11}c_{22}
\end{aligned}$$

são respectivamente o determinante e o negativo do determinante da matriz $[c_{ij}]$, cujo valor já foi demonstrado ser diferente de zero, o que implica na quarta propriedade das funções v_1 e v_2:

$$\begin{aligned}
\alpha_1 v_2(a) + \beta_1 v_2'(a) \neq 0 \\
\alpha_2 v_1(b) + \beta_2 v_1'(b) \neq 0
\end{aligned}$$

o que completa a demonstração. ■

E Operadores limitados

Teorema E.1. *Dado* $u \in \mathcal{C}(a,b)$ *com* $\|u\| = 1$ *e o operador* T *definido pela expressão* (279)*, a norma de* T *pode ser dada por*

$$\|T\| = \sup_{\|u\|=1} |\langle Tu, u\rangle|$$

Demonstração. Da definição (4.1), temos

$$\|T\| = \sup_{\|u\|=1} \|Tu\|$$

Mas também do teorema (1.10):

$$\|Tu\| = \sup_{\|w\|=1} |\langle w, Tu\rangle|$$

ou seja,

$$\|T\| = \sup_{\substack{\|u\|=1 \\ \|w\|=1}} |\langle w, Tu\rangle|$$

Uma vez que T é um operador hermitiano, temos

$$\langle Tu, w\rangle = \langle u, Tw\rangle \implies 2\langle w, Tu\rangle = \langle u, Tw\rangle + \langle Tu, w\rangle$$

Além disso,

$$\frac{1}{2}\left[\langle T(u+w), u+w\rangle - \langle T(w-u), w-u\rangle\right] =$$
$$\frac{1}{2}\left[\langle T(u+w), u\rangle + \langle T(u+w), w\rangle - \langle T(w-u), w\rangle + \langle T(w-u), u\rangle\right]$$

$$= \frac{1}{2}\left[\langle Tu, u\rangle + \langle Tw, u\rangle + \langle Tu, w\rangle + \langle Tw, w\rangle - \right.$$
$$\left.\langle Tw, w\rangle + \langle Tu, w\rangle + \langle Tw, u\rangle - \langle Tu, u\rangle\right]$$

$$\begin{aligned}
\frac{1}{2}\left[\langle T(u+w), u+w\rangle - \langle T(w-u), w-u\rangle\right] &= \frac{1}{2}\left[2\langle Tw, u\rangle + 2\langle Tu, w\rangle\right] \\
&= \langle Tw, u\rangle + \langle Tu, w\rangle \\
&= \overline{\langle u, Tw\rangle} + \langle Tu, w\rangle \\
&= \langle u, Tw\rangle + \langle Tu, w\rangle \\
&= 2\langle w, Tu\rangle
\end{aligned}$$

Por outro lado:

$$\begin{aligned}
\langle T(u+w), u+w\rangle &= \|u+w\|^2 \cdot \left\langle T\left(\frac{u+w}{\|u+w\|}\right), \frac{u+w}{\|u+w\|}\right\rangle \\
&\leq \|u+w\|^2 \cdot \sup_{\|u\|=1} |\langle Tu, u\rangle|
\end{aligned}$$

Analogamente, temos

$$\langle T(w-u), w-u\rangle \leq \|w-u\|^2 \cdot \sup_{\|u\|=1} |\langle Tu, u\rangle|$$

Isso implica

$$\frac{1}{2}\left[\langle T(u+w), u+w\rangle - \langle T(w-u), w-u\rangle\right] \leq$$
$$\frac{1}{2}\left[\left(\|u+w\|^2 + \|w-u\|^2\right) \cdot \sup_{\|u\|=1} |\langle Tu, u\rangle|\right]$$

Mas expandindo as normas de $u+w$ e $w-u$:

$$\begin{aligned}
\|u+w\|^2 + \|w-u\|^2 &= \langle u+w, u+w\rangle + \langle w-u, w-u\rangle \\
&= \langle u, u\rangle + 2\langle u, w\rangle + \langle w, w\rangle + \langle w, w\rangle - 2\langle u, w\rangle + \langle u, u\rangle \\
&= 2 \cdot [\|u\|^2 + \|w\|^2]
\end{aligned}$$

E então finalmente chegamos à expressão

$$2\langle w, Tu\rangle \leq \left(\|u\|^2 + \|w\|^2\right) \cdot \sup_{\|u\|=1} |\langle Tu, u\rangle|$$

Lembrando que $\|u\| = \|w\| = 1$, temos

$$2|\langle w, Tu\rangle| \leq (1+1) \cdot \sup_{\|u\|=1} |\langle Tu, u\rangle|$$

$$|\langle w, Tu\rangle| \leq \sup_{\|u\|=1} |\langle Tu, u\rangle|$$

Em particular:

$$\|T\| = \sup_{\substack{\|u\|=1 \\ \|w\|=1}} |\langle w, Tu\rangle| \leq \sup_{\|u\|=1} |\langle Tu, u\rangle| \tag{458}$$

Mas também pela desigualdade de Cauchy-Schwarz:

$$\sup_{\|u\|=1} |\langle Tu, u\rangle| \leq \sup_{\|u\|=1} \|Tu\| \cdot \|u\| \leq \|T\| \tag{459}$$

Juntando as inequações (458) e (459), concluímos que

$$\|T\| = \sup_{\|u\|=1} |\langle Tu, u\rangle|$$

■

F Autovalores da equação integral de Fredholm

Teorema F.1. *Ou $\|T\|$ ou $-\|T\|$ são autovalores de T.*

Demonstração. O teorema (4.17) garante que

$$\|T\| = \sup_{\|u\|=1} |\langle Tu, u\rangle|$$

Como as funções $T(x,s)$, $r(x)$ e $u(x)$ são reais, então $\langle Tu, u\rangle$ também é real. Daí segue que existem duas possibilidades:

$$\|T\| = \sup_{\|u\|=1} \langle Tu, u\rangle \quad \text{ou} \quad \|T\| = -\inf_{\|u\|=1} \langle Tu, u\rangle$$

Vamos supor o primeiro caso, uma vez que a demonstração do segundo é análogo. Como $\|T\|$ é o supremo de $\langle Tu, u\rangle$, com $\|u\| = 1$, podemos construir uma sequência $\{u_n\}$, $\|u_n\| = 1$, $u_n \in C(a,b)$ tal que

$$\langle Tu_n, u_n\rangle \to \|T\| \quad \text{quando} \quad n \to \infty$$

Dessa sequência, podemos construir uma outra, formada pela aplicação do operador T nos seus elementos: $\{Tu_n\}$, $\|u_n\| = 1$. O teorema (4.16) garante que

$\{Tu_n\}$ é uniformemente limitado e equicontínuo, e então podemos invocar o teorema de Ascolí-Arzela (4.18) nessa sequência para mostrar que existe uma subsequência $\{Tu_{nj}\}$ que converge uniformemente em $[a, b]$ para uma função contínua φ_0. Como consequência, temos

$$\forall x \in [a, b], \quad \left|Tu_{nj}(x) - \varphi_0(x)\right| \to 0 \quad \text{quando} \quad j \to \infty$$

ou seja,

$$\left\|Tu_{nj} - \varphi_0\right\| = \sqrt{\int_a^b \left|Tu_{nj}(x) - \varphi_0(x)\right|^2 r(x)\,\mathrm{d}x} \to 0$$

o que implica

$$\left\|Tu_{nj}\right\| - \|\varphi_0\| \leq \left\|Tu_{nj}(x) - \varphi_0(x)\right\| \to 0$$

$$\left\|Tu_{nj}\right\| \to \|\varphi_0\|$$

Agora calculando o quadrado da norma de $Tu_{nj} - \|T\|u_{nj}$:

$$\begin{aligned}
\left\|Tu_{nj} - \|T\|u_{nj}\right\|^2 &= \langle Tu_{nj} - \|T\|u_{nj}, Tu_{nj} - \|T\|u_{nj}\rangle \\
&= \left\|Tu_{nj}\right\|^2 - 2 \cdot \langle Tu_{nj}, \|T\|u_{nj}\rangle + \left\|\|T\|u_{nj}\right\|^2 \\
&= \left\|Tu_{nj}\right\|^2 - 2 \cdot \|T\| \langle Tu_{nj}, u_{nj}\rangle + \|T\|^2 \cdot \left\|u_{nj}\right\|^2 \qquad (460)
\end{aligned}$$

Uma vez que $\langle Tu_{nj}, u_{nj}\rangle \to \|T\|$ e $\left\|Tu_{nj}\right\|^2 \to \|\varphi_0\|^2$ quando $j \to \infty$, temos

$$\begin{aligned}
\left\|Tu_{nj} - \|T\|u_{nj}\right\|^2 &\to \|\varphi_0\|^2 - 2 \cdot \|T\|^2 + \|T\|^2 \\
&\to \|\varphi_0\|^2 - \|T\|^2
\end{aligned}$$

Mas como a norma $\left\|Tu_{nj} - \|T\|u_{nj}\right\|^2$ é sempre não negativa, segue que

$$\|\varphi_0\|^2 - \|T\|^2 \geq 0$$

$$\|\varphi_0\|^2 \geq \|T\|^2 > 0 \implies \|\varphi_0\|^2 > 0$$

uma vez que $\|T\| > 0$. Isso significa que $\|\varphi_0\|^2 \neq 0$ e $\varphi_0(x)$ não pode ser a função identicamente nula em $[a, b]$. Da inequação (284), temos

$$\left\|Tu_{nj}\right\|^2 \leq \|T\|^2 \cdot \left\|u_{nj}\right\|^2 = \|T\|^2$$

Portanto, da equação (460):

$$\begin{aligned}
\left\|Tu_{nj} - \|T\|u_{nj}\right\|^2 &\leq \|T\|^2 - 2 \cdot \|T\| \langle Tu_{nj}, u_{nj}\rangle + \|T\|^2 \\
&\leq 2 \cdot \|T\|^2 - 2 \cdot \|T\| \langle Tu_{nj}, u_{nj}\rangle
\end{aligned}$$

Lembrando que $\langle Tu_{nj}, u_{nj}\rangle \to \|T\|$ quando $j \to \infty$:

$$\left\|Tu_{nj} - \|T\|u_{nj}\right\|^2 \to 2 \cdot \|T\|^2 - 2 \cdot \|T\|^2 \to 0$$

$$\left\| Tu_{nj} - \|T\| u_{nj} \right\| \to 0$$

Por fim:

$$\begin{aligned} 0 \le \|T\varphi_0 - \|T\|\varphi_0\| \le& \\ &\left\| T\varphi_0 - T(Tu_{nj}) + T(Tu_{nj}) - \|T\|Tu_{nj} + \|T\|Tu_{nj} - \|T\|\varphi_0 \right\| \end{aligned}$$

$$\begin{aligned} 0 \le \|T\varphi_0 - \|T\|\varphi_0\| \le \left\| T\varphi_0 - T(Tu_{nj}) \right\| +& \\ &\left\| T(Tu_{nj}) - \|T\|Tu_{nj} \right\| + \left\| \|T\|Tu_{nj} - \|T\|\varphi_0 \right\| \end{aligned}$$

Quando $j \to \infty$ na primeira norma:

$$\begin{aligned} \left\| T\varphi_0 - T(Tu_{nj}) \right\| &= \left\| T\left[\varphi_0 - Tu_{nj}\right] \right\| \\ &\to 0 \quad \text{uma vez que} \quad Tu_{nj} \to \varphi_0 \end{aligned}$$

Para a segunda norma:

$$\begin{aligned} \left\| T(Tu_{nj}) - \|T\|Tu_{nj} \right\| &= \left\| T\left[Tu_{nj} - \|T\|u_{nj}\right] \right\| \\ &\le \|T\| \cdot \underbrace{\left\| Tu_{nj} - \|T\|u_{nj} \right\|}_{\to 0} \\ \left\| T(Tu_{nj}) - \|T\|Tu_{nj} \right\| &\to 0 \end{aligned}$$

e a terceira:

$$\begin{aligned} \left\| \|T\|Tu_{nj} - \|T\|\varphi_0 \right\| &= \|T\| \cdot \underbrace{\left\| Tu_{nj} - \varphi_0 \right\|}_{\to 0} \\ \left\| \|T\|Tu_{nj} - \|T\|\varphi_0 \right\| &\to 0 \end{aligned}$$

Ou seja, o valor constante $\|T\varphi_0 - \|T\|\varphi_0\|$ é confrontado inferiormente por 0 e superiormente por uma sequência que tende a zero. Isso implica

$$\begin{aligned} \|T\varphi_0 - \|T\|\varphi_0\| = 0 \implies T\varphi_0 - \|T\|\varphi_0 = 0 \\ T\varphi_0(x) = \|T\|\varphi_0(x) \end{aligned} \tag{461}$$

Como φ_0 é contínua em $[a, b]$, a equação (461) é válida para $\forall x \in [a, b]$. Em outras palavras, $\|T\|$ é um autovalor de T cuja autofunção é $\varphi_0(x)$.

E no caso alternativo em que

$$\|T\| = -\inf_{\|u\|=1} \langle Tu, u \rangle$$

podemos construir uma sequência $\{u_n\}$, $\|u_n\| = 1$, $u_n \in C(a, b)$ tal que

$$\langle Tu_n, u_n \rangle \to -\|T\| \quad \text{quando} \quad n \to \infty$$

Refazendo os passos acima, isso leva à conclusão que $-\|T\|$ é um autovalor de T.

■

G Polinômios de Hermite

Teorema G.1. *Dado $\mathcal{L}^2$ o espaço euclidiano das funções quadrado integráveis em todo o $\mathbb{R}$ com o produto interno definido por*

$$\langle f, g\rangle = \int_{-\infty}^{\infty} f(s)g(s)\,\mathrm{d}s$$

segue que

$$\begin{aligned}\left\|H_n(\xi)e^{-\xi^2/2}\right\|^2 &= \int_{-\infty}^{\infty} H_n(s)e^{-s^2/2}H_n(s)e^{-s^2/2}\,\mathrm{d}s \\ &= \int_{-\infty}^{\infty} H_n^2(s)e^{-s^2}\,\mathrm{d}s \\ &= 2^n n!\sqrt{\pi}\end{aligned} \tag{462}$$

onde H_n é o polinômio de Hermite de grau n cuja potência mais alta tem um fator de 2^n.

Demonstração. Os polinômios de Hermite também podem ser definidos pela fórmula de Rodrigues:

$$H_n(s) = (-1)^n e^{s^2}\frac{d^n}{ds^n}e^{-s^2} \tag{463}$$

Substituindo (463) em (462), obtemos

$$\begin{aligned}\int_{-\infty}^{\infty} H_n^2(s)e^{-s^2}\,\mathrm{d}s &= \int_{-\infty}^{\infty} H_n(s)H_n(s)e^{-s^2}\,\mathrm{d}s \\ &= (-1)^n\int_{-\infty}^{\infty}\left(e^{s^2}\frac{d^n}{ds^n}e^{-s^2}\right)H_n(s)e^{-s^2}\,\mathrm{d}s \\ &= (-1)^n\int_{-\infty}^{\infty} H_n(s)\frac{d^n}{ds^n}e^{-s^2}\,\mathrm{d}s\end{aligned}$$

Integrando a última expressão por partes, onde a exponencial é integrada e o polinômio é derivado, temos

$$\begin{aligned}\int_{-\infty}^{\infty} H_n^2(s)e^{-s^2}\,\mathrm{d}s &= (-1)^n H_n(s)\frac{d^{n-1}}{ds^{n-1}}e^{-s^2}\Bigg|_{-\infty}^{\infty} - (-1)^n\int_{-\infty}^{\infty}\frac{d}{ds}H_n(s)\frac{d^{n-1}}{ds^{n-1}}e^{-s^2}\,\mathrm{d}s \\ &= (-1)^{n+1}\int_{-\infty}^{\infty}\frac{d}{ds}H_n(s)\frac{d^{n-1}}{ds^{n-1}}e^{-s^2}\,\mathrm{d}s\end{aligned}$$

O primeiro termo da primeira linha desaparece uma vez que a derivada $n-1$ da exponencial é a própria exponencial vezes um polinômio, e esse termo decai rapidamente para zero em $-\infty$ e ∞.

Repetindo a integração por parte outras $n-1$ vezes, obteremos

$$\int_{-\infty}^{\infty} H_n^2(s)e^{-s^2}\,\mathrm{d}s = (-1)^{n+n}\int_{-\infty}^{\infty} e^{-s^2}\frac{d^n}{ds^n}H_n(s)\,\mathrm{d}s$$

A n-ésima derivada de H_n é $2^n n!$. Para mostrar isso, basta ver que H_n é um polinômio de grau n, e então apenas o termo de maior grau, n, não se anulará após as n derivadas. Como 2^n é o coeficiente da potência mais alta e a n-ésima derivada de s^n é igual a $n!$, temos que

$$\frac{d^n}{ds^n} H_n(s) = 2^n n!$$

e portanto

$$\begin{aligned}\int_{-\infty}^{\infty} H_n^2(s) e^{-s^2} \,\mathrm{d}s &= (-1)^{2n} 2^n n! \int_{-\infty}^{\infty} e^{-s^2} \,\mathrm{d}s \\ &= 2^n n! \sqrt{\pi}\end{aligned}$$

o que completa a demonstração. ■

Teorema G.2.

$$\int_{-\infty}^{\infty} s^k H_n(s) e^{-s^2} \,\mathrm{d}s = 0 \quad \textit{se} \quad k < n$$

Demonstração. Substituindo H_n pela expressão (463), temos

$$\begin{aligned}\int_{-\infty}^{\infty} s^k H_n(s) e^{-s^2} \,\mathrm{d}s &= (-1)^n \int_{-\infty}^{\infty} \left(e^{s^2} \frac{d^n}{ds^n} e^{-s^2} \right) s^k e^{-s^2} \,\mathrm{d}s \\ &= (-1)^n \int_{-\infty}^{\infty} s^k \frac{d^n}{ds^n} e^{-s^2} \,\mathrm{d}s\end{aligned}$$

Realizando a integração por partes, onde derivamos s^k e integramos a n-ésima derivada da exponencial, obtemos

$$\begin{aligned}\int_{-\infty}^{\infty} s^k H_n(s) e^{-s^2} \,\mathrm{d}s &= (-1)^n s^k \frac{d^{n-1}}{ds^{n-1}} e^{-s^2} \Bigg|_{-\infty}^{\infty} - (-1)^n \int_{-\infty}^{\infty} \frac{d}{ds} s^k \frac{d^{n-1}}{ds^{n-1}} e^{-s^2} \,\mathrm{d}s \\ &= (-1)^{n+1} \int_{-\infty}^{\infty} k s^{k-1} \frac{d^{n-1}}{ds^{n-1}} e^{-s^2} \,\mathrm{d}s\end{aligned}$$

O primeiro termo da primeira linha desaparece uma vez que a derivada $n-1$ da exponencial é a própria exponencial vezes um polinômio, e esse termo decai rapidamente para zero em $-\infty$ e ∞.

Repetindo o procedimento mais $k-1$ vezes, obtemos

$$\begin{aligned}\int_{-\infty}^{\infty} s^k H_n(s) e^{-s^2} \,\mathrm{d}s &= (-1)^{n+k} \int_{-\infty}^{\infty} \frac{d^k}{ds^k} s^k \frac{d^{n-k}}{ds^{n-k}} e^{-s^2} \,\mathrm{d}s \\ &= (-1)^{n+k} \int_{-\infty}^{\infty} k! \frac{d^{n-k}}{ds^{n-k}} e^{-s^2} \,\mathrm{d}s\end{aligned}$$

Podemos repetir a integração por partes mais uma vez, mas nesse caso a derivada de $k!$ é nula e então

$$\begin{aligned}\int_{-\infty}^{\infty} s^k H_n(s) e^{-s^2}\,\mathrm{d}s &= (-1)^{n+k} k! \frac{d^{n-k-1}}{ds^{n-k-1}} e^{-s^2}\Bigg|_{-\infty}^{\infty} \\ &= 0\end{aligned}$$

o que completa a demonstração. ■

H Funções de Bessel esféricas

H.1 Códigos em Python

H.1.1 Encontrar as raízes e os coeficientes

```python
import numpy as np
import math
import os

N_MAX = 200             # número de coeficientes da série de taylor (não recomendo aumentar muito, pois pode dar o
L_MAX = 6               # número + 1 de funções de Bessel esféricas para serem calculadas (l = 0 também é calculad
R_MAX = 9               # número de raízes para serem encontradas
delta = 10**(-1)        # precisão inicial para encontrar mudança do sinal
precisao = 10**(-10)        # precisão de casa decimal para encontrar raízes

# OBSERVAÇÃO: quando mais distante a raíz estiver da origem, maior será o erro
# devido a soma finita da série de Taylor.
# Normalmente erros na primeira casa decimal aparecem aproximadamente quando r > 35

A = np.zeros((N_MAX,L_MAX+1))   # matriz dos coeficientes para cada l
salvarcoef = True
salvarraiz = True

dirname = os.path.dirname(__file__)
filenamecoef = os.path.join(dirname, 'besselcoef.txt')
filenameroot = os.path.join(dirname, 'besselroot.txt')

def calcularcoef():         # usar a relação de recorrência e guardar os resultados no vetor A
    for li in range(L_MAX+1):
        A[0,li] = 1
        A[1,li] = 0
        for n in range(N_MAX-2):
            if (n % 2 == 0):
                A[n+2,li] = (-1)*A[n,li]/((n+2)*(n+3+(2*li)))       # relação de recorrencia

    if(salvarcoef):
```

```
32          # salvar os dados em um arquivo de texto
33          coef_file = open(filenamecoef, "w")
34          for row in A:
35              np.savetxt(coef_file, row)
36          coef_file.close()
37      return 1
38
39  def funcao(r,l):                                # função esférica de Bessel
40      i = N_MAX - 1
41      soma = 0
42      while(i >= 0):
43          soma = soma*r + A[i,l]
44          i = i - 1
45      return math.pow(r,l)*soma
46
47  def calcularraiz():         # calcular raízes usando variáveis antes e depois para comparar o que aconteceu entre
48      B = np.zeros((R_MAX,L_MAX+1))           # matriz das raízes para cada l
49      for li in range(L_MAX+1):
50          x = 0
51          b = funcao(0,li)
52          bx = 0
53          i = 0
54          raiz = -1
55          while(i < R_MAX):             # parar apenas quando atingir o número de raízes requisitadas
56              a = funcao(x,li)
57              ax = x
58              if(a*b < 0):              # inverteu o sinal, há uma raíz entre os pontos
59                  centro = -((a/(b-a))*(bx-ax)) + ax
60                  a = funcao(ax,li)
61                  b = funcao(bx,li)
62                  while(True):
63                      if(abs(a) < abs(b)):  # qual extremo está mais perto da raiz
64                          bx = centro
65                      else:
66                          ax = centro
67
68                      if(b-a == 0):   # evitar divisão por zero se a subtração for muito pequena
69                          B[i,li] = centro
70                          break
71                      centro = -((a/(b-a))*(bx-ax)) + ax
72
73                      if(abs(funcao(centro,li)) < precisao):  # quando a função é menor do que a precisão
74                          B[i,li] = centro      # raiz
75                          break                 # finalizar busca dessa raiz específica, procurar outras
76                      a = funcao(ax,li)         # atualizar valores dos extremos
77                      b = funcao(bx,li)
78                  print("Raíz ",i+1," / L =", li, ": ", B[i,li])
79                  i += 1
```

```
80              b = funcao(x,li)
81              bx = x
82              x += delta
83
84      if(salvarraiz):
85          # salvar os dados em um arquivo de texto
86          root_file = open(filenameroot, "w")
87          for row in B:
88              np.savetxt(root_file, row)
89          root_file.close()
90      return 1
91
92  def efetuar(coef,rooot):
93      if(coef == 1):
94          print("Calculando coeficientes...")
95          calcularcoef()
96      if(rooot == 1):
97          print("Calculando raízes...")
98          calcularraiz()
99      print("Pronto!")
100
101 # calcular coeficientes ou raízes ou ambos
102 cof = 1
103 rot = 1
104
105 # se não calcular novos coeficientes, ler a matriz dos coeficientes já calculados anteriormente
106 if(cof == 0):
107     A = np.loadtxt(filenamecoef).reshape(N_MAX, L_MAX+1)
108 efetuar(cof,rot)     # calcular coeficientes ou raízes
```

H.1.2 Encontrar as normas

```
1  import numpy as np
2  import math
3  import os
4
5  # nota:
6  # essas três constantes precisam ser os mesmos do código que gera os coeficientes
7  # e as que são usadas na expansão em autofunções
8  N_MAX = 200
9  L_MAX = 6
10 R_MAX = 9
11
12 dirname = os.path.dirname(__file__)
13 filenamecoef = os.path.join(dirname, 'besselcoef.txt')
14 filenameroot = os.path.join(dirname, 'besselroot.txt')
```

```
15  filenamenorm = os.path.join(dirname, 'besselnorm.txt')
16  filenamenorm2 = os.path.join(dirname, 'besselnorm2.txt')
17
18  A = np.loadtxt(filenamecoef).reshape(N_MAX, L_MAX+1)     # ler matriz de coeficientes
19  B = np.loadtxt(filenameroot).reshape(R_MAX, L_MAX+1)    # ler matriz de raízes
20  C = np.zeros((R_MAX,L_MAX+1))   # matriz onde ficarão guardados as normas em relação a um raio especifico
21  D = np.zeros((L_MAX+1))   # matriz onde ficarão guardados as normas em relação ao infinito
22
23  r00 = 3                     # raio da esfera (precisa ser o mesmo em todos os códigos)
24  delta = 10**(-2)            # tamanho da partição da integral
25  delta2 = 10**(-2)           # tamanho da partição da integral no caso infinito
26
27  def funcao(r,l):                                  # função esférica de Bessel
28      i = N_MAX - 1
29      soma = 0
30      while(i >= 0):
31          soma = soma*r + A[i,l]
32          i = i - 1
33      return math.pow(r,l)*soma
34
35  def funcao2(r,l,k):                               # integrando da norma
36      u = funcao(k*r,l)*r
37      return u*u
38
39  def gerarnormas(r0):        # calcular as normas usando regra de Simpson (normas das autofunções)
40      for li in range(L_MAX+1):
41          print((li/(L_MAX+1))*100," %")      # progresso
42          for rt in range(R_MAX):
43              integral = 0
44              k = B[rt,li]/r0
45              x = 0
46              while(x <= r0):
47                  integral += (delta/6)*(funcao2(x,li,k) + 4*funcao2(x+(0.5*delta),li,k) + funcao2(x+delta,li,k))
48                  x += delta
49              C[rt,li] = math.sqrt(integral)
50      # salvar os dados em um arquivo de texto
51      norm_file = open(filenamenorm, "w")
52      for row in C:
53          np.savetxt(norm_file, row)
54      norm_file.close()
55
56  def gerarnormas2():  # calcular as normas usando regra de Simpson (normas do intervalo infinito)
57      for li in range(L_MAX+1):
58          print((li/(L_MAX+1))*100," %")      # progresso
59          integral = 0
60          r0 = 18          # número grande
61          k = 1
62          x = 0
```

```
            while(x <= r0):
                integral += (delta/6)*(funcao2(x,li,k) + 4*funcao2(x+(0.5*delta),li,k) + funcao2(x+delta,li,k))
                x += delta2
            D[li] = math.sqrt(integral)
        # salvar os dados em um arquivo de texto
        norm_file2 = open(filenamenorm2, "w")
        np.savetxt(norm_file2, D)
        norm_file2.close()
        return 1

    print("Gerando normas...")
    gerarnormas(r00)    # normas para as autofunções
    #gerarnormas2()       # normas para intervalo infinito
    print("Pronto!")
```

H.1.3 Gráfico das funções

```
import numpy as np
import math
import matplotlib.pyplot
import os

N_MAX = 200
L_MAX = 6
R_MAX = 9

dirname = os.path.dirname(__file__)
filenamecoef = os.path.join(dirname, 'besselcoef.txt')
filenamenorm2 = os.path.join(dirname, 'besselnorm2.txt')

A = np.loadtxt(filenamecoef).reshape(N_MAX, L_MAX+1)    # ler matriz dos coeficientes
D = np.loadtxt(filenamenorm2).reshape(L_MAX+1)    # ler matriz das normas

lmax = 2            # plotar até qual l
linit = 0           # plotar a partir de qual l
x = 0               # começo
xf = 20             # fim
npartes = 3000      # quantidade de pontos a serem calculadoss

def funcao(r,l):                             # função esférica de Bessel
    i = N_MAX - 1
    soma = 0
    while(i >= 0):
        soma = soma*r + A[i,l]
        i = i - 1
    return math.pow(r,l)*soma
```

```
30
31  ex = []
32  ey = []
33  ex1 = []
34  ey1 = []
35
36  delta = (xf - x)/npartes
37  ex1.append(0)
38  ex1.append(xf)
39  ey1.append(0)
40  ey1.append(0)
41
42  matrizvariavel = np.linspace(x,xf,npartes)
43  matrizcalculos = np.zeros((len(matrizvariavel),L_MAX+1))
44
45  for li in range(linit,lmax+1,1):
46      for i in range(len(matrizvariavel)):
47          matrizcalculos[i,li] = funcao(matrizvariavel[i],li)/D[li]
48
49  l = linit
50  while(l <= lmax):        # plotar cada gráfico
51      matplotlib.pyplot.plot(matrizvariavel, matrizcalculos[:,l],label='l = ' + str(l))
52      l += 1
53
54  matplotlib.pyplot.plot(ex1, ey1)
55  matplotlib.pyplot.title(label='Funções de Bessel esféricas normalizadas')
56  matplotlib.pyplot.xlabel("r")
57  matplotlib.pyplot.ylabel("j(r)")
58  matplotlib.pyplot.legend()
59  matplotlib.pyplot.show()
```

H.1.4 Expansão em autofunções de Sturm-Liouville

```
1   import numpy as np
2   import math
3   import matplotlib.pyplot
4   import os
5
6   # nota:
7   # essas três constantes precisam ser os mesmos do código que gera os coeficientes
8   # e as que são usadas na expansão em autofungções
9   N_MAX = 200
10  L_MAX = 6
11  R_MAX = 9
12
13  dirname = os.path.dirname(__file__)
```

```
14  filenamecoef = os.path.join(dirname, 'besselcoef.txt')
15  filenameraiz = os.path.join(dirname, 'besselroot.txt')
16  filenamenorm = os.path.join(dirname, 'besselnorm.txt')
17
18  A = np.loadtxt(filenamecoef).reshape(N_MAX, L_MAX+1)    # ler matriz de coeficientes
19  B = np.loadtxt(filenameraiz).reshape(R_MAX, L_MAX+1)   # ler matriz de raízes
20  C = np.loadtxt(filenamenorm).reshape(R_MAX, L_MAX+1)    # ler matriz das normas
21
22  D = np.zeros(R_MAX)     # vetor dos coeficientes da expansão em autofunções
23  l = 0                   # escolha um l
24  delta = 10**(-2)        # precisão da integral
25  r0 = 3                  # raio da esfera de vácuo
26
27  ncoeficientes = 9     # número de coeficientes da expansão em autofunções
28
29  if(ncoeficientes > D.size):
30      raise Exception("Não há raízes suficientes!")
31
32  def funcao(r,l):                                # função esférica de Bessel
33      i = N_MAX - 1
34      soma = 0
35      while(i >= 0):
36          soma = soma*r + A[i,l]
37          i = i - 1
38      return math.pow(r,l)*soma
39
40  def g(x):                                       # qualquer função quadrado-integrável em [0,r0]
41      return math.pow(math.e,0-math.pow((x-(1/5)*r0),2))
42
43  def funcintegra(r,k1,rt,l):                     # usado na integral dos coeficientes da expansão
44      return g(r)*funcao(k1*r,l)*r*r/C[rt,l]
45
46  def calcularcoord():                            # calcular coeficientes da expansão usando regra de Simpson
47      for rt in range(R_MAX):
48          k = B[rt,l]/r0
49          integral = 0
50          x = 0
51          while(x <= r0):
52              integral += (delta/6)*(funcintegra(x,k,rt,l) + 4*funcintegra(x+(0.5*delta),k,rt,l) + funcintegra(x+de
53              x += delta
54          D[rt] = integral
55
56  ex = []
57  ey = []
58  ex1 = []
59  ey1 = []
60  ey2 = []
61
```

```python
# parâmetros para imprimir o gráfico
x = 0                      # inicio
xf = 3                     # fim (geralmente é r0, pois a convergência não é assegurada para r > r0)
npartes = 1000             # número de partições
delta2 = (xf-x)/npartes    # tamanho das partições

# colocar o eixo x
ex1.append(0)
ex1.append(xf)
ey1.append(0)
ey1.append(0)

def expansao(r):          # calcular a expansão em si
    soma = 0
    for rt in range(ncoeficientes):
        k = B[rt,l]/r0                          # para cada raíz 'rt', um 'k' diferente
        soma += D[rt]*funcao(k*r,l)/C[rt,l]     # o somatório da expansão
    return soma

print("Calculando coordenadas...")
calcularcoord()
print("Calculando a expansão...")

# calcular a expansão para cada 'x' para imprimir o gráfico
while(x <= xf):
    ex.append(x)
    ey.append(expansao(x))
    ey2.append(g(x))
    x += delta2
print("Pronto!")

matplotlib.pyplot.plot(ex, ey, label='exp') # label='l = 0'
matplotlib.pyplot.plot(ex1, ey1)
matplotlib.pyplot.plot(ex, ey2, label='g(r)')
matplotlib.pyplot.title(label='Expansão em autofunções de Bessel esféricas com l = ' + str(l))
matplotlib.pyplot.xlabel("r")
matplotlib.pyplot.ylabel("g(r)")
matplotlib.pyplot.legend()
matplotlib.pyplot.show()
```

I Polinômios de Legendre

Sejam $P(x)$ as funções que satisfazem a equação diferencial de Legendre

$$(1-x^2)P''(x) - 2xP'(x) + \lambda P(x) = 0 \qquad (464)$$

ou equivalentemente

$$P''(x) - x^2 P''(x) - 2xP'(x) + \lambda P(x) = 0 \tag{465}$$

As funções $1 - x^2$ e $-2x$ em (464) são analíticas e portanto podemos resolvê-la usando o método das séries de potências centradas em $x = 0$. Mas como $1 - x^2$ se anula em $x = \pm 1$, então a equação possui singularidades em $x = \pm 1$ e consequentemente a convergência da solução obtida pelo método de série de potências estará restrita ao intervalo $(-1, 1)$. Fazendo

$$P(x) = \sum_{n=0}^{\infty} a_n x^n \tag{466}$$

então

$$P'(x) = \sum_{n=0}^{\infty} n a_n x^{n-1}$$

$$P''(x) = \sum_{n=0}^{\infty} n(n-1) a_n x^{n-2} = \sum_{n=2}^{\infty} n(n-1) a_n x^{n-2} = \sum_{n=0}^{\infty} (n+2)(n+1) a_{n+2} x^n$$

Inserindo essas expressões em (465), obtemos

$$\sum_{n=0}^{\infty} (n+2)(n+1) a_{n+2} x^n - x^2 \sum_{n=0}^{\infty} n(n-1) a_n x^{n-2} - 2x \sum_{n=0}^{\infty} n a_n x^{n-1} + \lambda \sum_{n=0}^{\infty} a_n x^n = 0$$

$$\sum_{n=0}^{\infty} \left[(n+2)(n+1) a_{n+2} - n(n-1) a_n - 2n a_n + \lambda a_n \right] x^n = 0$$

Devido à independência linear dos monômios x^n, temos que $\forall n \geq 0$,

$$(n+2)(n+1) a_{n+2} - n(n-1) a_n - 2n a_n + \lambda a_n = 0$$

$$\begin{aligned} (n+2)(n+1) a_{n+2} &= \left[n(n-1) + 2n - \lambda \right] a_n \\ &= \left[n(n+1) - \lambda \right] a_n \end{aligned}$$

e portanto a relação de recorrência é

$$a_{n+2} = \frac{\left[n(n+1) - \lambda \right] a_n}{(n+2)(n+1)} \quad \forall n \geq 0 \tag{467}$$

com a_0 e a_1 arbitrários. Isso significa que os coeficientes pares dependem do valor de a_0 enquanto que os ímpares dependem de a_1. Então podemos escrever a solução na forma

$$P(x) = \sum_{n=0}^{\infty} a_{2n} x^{2n} + \sum_{n=0}^{\infty} a_{2n+1} x^{2n+1} \tag{468}$$

Se n é muito grande, podemos desprezar λ, no que resulta na aproximação

$$a_{n+2} \approx \frac{n(n+1)a_n}{(n+2)(n+1)} \approx \frac{n}{n+2}a_n$$

Para os coeficientes ímpares, temos

$$a_{2n+3} \approx \frac{2n+1}{2n+3}a_{2n+1} \tag{469}$$

Os primeiros termos são

$$a_3 = \frac{a_1}{3} \qquad a_5 = \frac{3}{5}a_1 = \frac{a_1}{5}$$

$$a_7 = \frac{5}{7}a_5 = \frac{a_1}{7} \qquad a_9 = \frac{a_1}{9}$$

Isso nos sugere que

$$a_{2n+1} = \frac{a_1}{2n+1} \tag{470}$$

Usando a relação de recorrência (469), temos que

$$a_{2n+3} = \frac{2n+1}{2n+3}a_{2n+1} = \frac{2n+1}{2n+3} \cdot \frac{a_1}{2n+1} = \frac{a_1}{2n+3}$$

o que prova (470) por indução.

Fazendo $a_0 = 0$ e $a_1 = 1$, obtemos

$$P(x) \approx \sum_{n=0}^{\infty} \frac{x^{2n+1}}{2n+1}$$

Agora vamos verificar o comportamento dos coeficientes ímpares de $P(x)$ quando $x \to 1$. Temos que

$$\lim_{x\to 1} P(x) = \sum_{n=0}^{\infty} \frac{1}{2n+1}$$

Uma vez que $f(x) = \dfrac{1}{2x+1}$ é sempre positiva em $x > 0$ e ela possui uma derivada sempre negativa nesse intervalo, podemos usar o teste da integral (2.8) para verificar se a série $\sum 1/(2n+1)$ converge ou não. Temos que

$$\lim_{k\to\infty} \int_1^k \frac{\mathrm{d}x}{2x+1} = \frac{1}{2}\lim_{k\to\infty} \ln(2k+1) - \frac{1}{2}\ln(3)$$

que diverge. Ou seja, $P(x)$ diverge em $x = 1$.

Para os coeficientes pares, temos

$$a_{2n+2} \approx \frac{2n}{2n+2}a_{2n} = \frac{n}{n+1}a_{2n} \tag{471}$$

Como essa é uma aproximação, vamos considerar que $a_2 \neq 0$, e então os próximos termos serão

$$a_4 = \frac{1}{2}a_2 \quad a_6 = \frac{2}{3}a_4 = \frac{2}{3 \cdot 2}a_2 = \frac{a_2}{3}$$

$$a_8 = \frac{a_2}{4} \quad a_{10} = \frac{a_2}{10}$$

o que sugere

$$a_{2n} \approx \frac{a_2}{n}$$

Substituindo essa expressão na relação (471), vemos que ela é satisfeita para todos os termos. E então, para a parte par de $P(x)$ temos

$$\lim_{x \to \pm 1} P(x) \approx \sum_{n=0}^{\infty} \frac{a_2}{n}(\pm 1)^{2n} = \sum_{n=0} \infty \frac{a_2}{n}$$

que também diverge, conforme vimos no exemplo (2.3). Ou seja, $P(x)$ diverge tanto em $x = 1$ quanto em $x = -1$.

Se quisermos evitar essa divergência em $x = \pm 1$, será preciso escolher λ de forma que trunque pelo menos uma das séries de (468) para torná-la uma soma finita. Ao verificar a relação de recorrência (467), vemos que $\lambda = l(l + 1)$ para um l inteiro não-negativo é suficiente para tal. Se l for par, a série das potências pares será finita, mas a série das potências ímpares continuará divergindo. Para evitar isso devemos escolher $a_1 = 0$ para anular a série das potências ímpares. Analogamente, se l for ímpar, a série de potências ímpares será finita e devemos escolher $a_0 = 0$. As soluções serão, portanto, polinômios. Esses são denominados de polinômios de Legendre.

Como

$$a_{n+2} = \frac{[n(n+1) - l(l+1)]\, a_n}{(n+2)(n+1)} \quad \forall n \geq 0 \tag{472}$$

segue que

$$a_{l+2} = \frac{[l(l+1) - l(l+1)]\, a_l}{(l+2)(l+1)} = 0$$

e portanto a potência não-nula mais alta do polinômio será l, o que implica em um polinômio de grau l.

Além disso, devido as imposições $a_1 = 0$ se l é par e $a_0 = 0$ se l for ímpar, então a equação

$$P(x) = \sum_{n=0}^{\infty} a_{2n}x^{2n} + \sum_{n=0}^{\infty} a_{2n+1}x^{2n+1} \tag{473}$$

nos diz que se l for par, o polinômio de Legendre de grau l só terá potências pares, o que implica em uma função par. Analogamente, se l for ímpar, o polinômio será ímpar.

I.1 Normalização das funções associadas

Teorema I.1. *Sejam as funções $P_l^m(x)$ dadas por*

$$P_l^m(x) = (-1)^m \frac{1}{2^l l!}(1-x^2)^{m/2} \frac{d^{l+m}}{dx^{l+m}}(1-x^2)^l \qquad l \geq m$$

Então

$$\int_{-1}^{1} P_l^m(x) P_l^m(x)\,\mathrm{d}x = \int_{-1}^{1} \left[P_l^m(x)\right]^2 \mathrm{d}x = \frac{2}{2l+1} \cdot \frac{(l+m)!}{(l-m)!} = A_{lm}$$

Demonstração. O quadrado de $P_l^m(x)$ é dado por

$$\left[P_l^m(x)\right]^2 = (-1)^{2m} \frac{1}{2^{2l}(l!)^2}(1-x^2)^m \frac{d^{l+m}}{dx^{l+m}}(1-x^2)^l \frac{d^{l+m}}{dx^{l+m}}(1-x^2)^l$$

e portanto

$$\int_{-1}^{1} \left[P_l^m(x)\right]^2 \mathrm{d}x = \frac{1}{2^{2l}(l!)^2} \int_{-1}^{1} (1-x^2)^m \frac{d^{l+m}}{dx^{l+m}}(1-x^2)^l \frac{d^{l+m}}{dx^{l+m}}(1-x^2)^l \,\mathrm{d}x$$

Iremos realizar uma integração por partes $l+m$ vezes, integrando o termo $\frac{d^{l+m}}{dx^{l+m}}(1-x^2)^l$ e derivando o restante. O termo que aparecerá fora de qualquer integral terá um fator $(1-x^2)^m$, e então sempre será zero, uma vez que os limites de integração são -1 e 1. Logo

$$A_{lm} = \frac{(-1)^{l+m}}{2^{2l}(l!)^2} \int_{-1}^{1} (1-x^2)^l \frac{d^{l+m}}{dx^{l+m}} \left[(1-x^2)^m \frac{d^{l+m}}{dx^{l+m}}(1-x^2)^l\right] \mathrm{d}x$$

onde o fator $(-1)^{l+m}$ veio da integração por partes aplicada $l+m$ vezes.

Agora vamos usar a fórmula de Leibniz para a derivada do produto,

$$\frac{d^{l+m}}{dx^{l+m}} \left[(1-x^2)^m \frac{d^{l+m}}{dx^{l+m}}(1-x^2)^l\right] = \sum_{k=0}^{l+m} \binom{l+m}{k} \frac{d^k}{dx^k}(1-x^2)^m \frac{d^{2l+2m-k}}{dx^{2l+2m-k}}(1-x^2)^l$$

Sabemos que $(1-x^2)^l$ é um polinômio de grau $2l$, enquanto que $(1-x^2)^m$ tem grau $2m$. Isso significa que se $k > 2m$, a primeira derivada será nula. E se $2l + 2m - k > 2l$, a segunda derivada será nula. Então segue que os únicos termos não nulos serão aqueles que k satisfaz $2m \leq k \leq 2m$, ou seja, um único termo em que $k = 2m$. Então a derivada do produto simplifica para

$$\begin{aligned}\frac{d^{l+m}}{dx^{l+m}} \left[(1-x^2)^m \frac{d^{l+m}}{dx^{l+m}}(1-x^2)^l\right] &= \binom{l+m}{2m} \frac{d^{2m}}{dx^{2m}}(1-x^2)^m \frac{d^{2l+2m-2m}}{dx^{2l+2m-2m}}(1-x^2)^l \\ &= \binom{l+m}{2m} \frac{d^{2m}}{dx^{2m}}(1-x^2)^m \frac{d^{2l}}{dx^{2l}}(1-x^2)^l\end{aligned}$$

Para encontrarmos os valores dessas derivadas de ordem $2m$ e $2l$, precisamos expandir as potências de $(1-x^2)$ usando o teorema binomial,

$$(1-x^2)^m = \sum_{r=0}^{m} \binom{m}{r} (-x^2)^{m-r} = \sum_{r=0}^{m} (-1)^{m-r} \binom{m}{r} x^{2(m-r)}$$

e então

$$\frac{d^{2m}}{dx^{2m}}(1-x^2)^m = \sum_{r=0}^{m} (-1)^{m-r} \binom{m}{r} \frac{d^{2m}}{dx^{2m}} x^{2(m-r)}$$

O único termo que não será nulo será quando $r=0$, ou seja,

$$\begin{aligned} \frac{d^{2m}}{dx^{2m}}(1-x^2)^m &= (-1)^m \binom{m}{0} \frac{d^{2m}}{dx^{2m}} x^{2m} \\ &= (-1)^m (2m)! \end{aligned}$$

De maneira totalmente análoga,

$$\frac{d^{2l}}{dx^{2l}}(1-x^2)^l = (-1)^l (2l)!$$

Substituindo todas essas expressões na integral e sabendo que

$$\binom{l+m}{2m} = \frac{(l+m)!}{(2m)!(l+m-2m)!} = \frac{(l+m)!}{(2m)!(l-m)!}$$

iremos obter

$$\begin{aligned} A_{lm} &= \frac{(-1)^{l+m}(-1)^{l+m}}{2^{2l}(l!)^2}(2l)!(2m)!\frac{(l+m)!}{(2m)!(l-m)!}\int_{-1}^{1}(1-x^2)^l\,\mathrm{d}x \\ &= \frac{(2l)!}{2^{2l}(l!)^2}\frac{(l+m)!}{(l-m)!}\int_{-1}^{1}(1-x^2)^l\,\mathrm{d}x \end{aligned} \tag{474}$$

Para calcular a integral que resta, iremos fazer a substituição $x = \cos(\theta)$, $\mathrm{d}x = -\sin(\theta)\,\mathrm{d}\theta$, de forma que

$$(1-x^2)^l = (1-\cos^2(\theta))^l = (\sin^2(\theta))^l = (\sin(\theta))^{2l}$$

$$x \to -1 \implies \theta \to \pi \quad \text{e} \quad x \to 1 \implies \theta \to 0$$

e então

$$\int_{-1}^{1}(1-x^2)^l\,\mathrm{d}x = \int_{\pi}^{0} -(\sin(\theta))^{2l}\sin(\theta)\,\mathrm{d}\theta = \int_{0}^{\pi}(\sin(\theta))^{2l+1}\,\mathrm{d}\theta$$

Agora precisamos usar o fato de que

$$\begin{aligned} \frac{d}{d\theta}\left(\sin^{n-1}(\theta)\cos(\theta)\right) &= (n-1)\sin^{n-2}(\theta)\cos^2(\theta) - \sin^n(\theta) \\ &= (n-1)\sin^{n-2}(\theta)(1-\sin^2(\theta)) - \sin^n(\theta) \\ &= (n-1)\sin^{n-2}(\theta) - (n-1)\sin^n(\theta) - \sin^n(\theta) \\ &= (n-1)\sin^{n-2}(\theta) - n\sin^n(\theta) \end{aligned}$$

e portanto

$$\sin^n(\theta) = \frac{(n-1)}{n}\sin^{n-2}(\theta) - \frac{1}{n}\frac{d}{d\theta}\left(\sin^{n-1}(\theta)\cos(\theta)\right)$$

Aplicando a integral em ambos os lados em um intervalo $0 \leq \theta \leq \pi$, obteremos

$$\begin{aligned}
\int_0^\pi \sin^n(\theta)\,\mathrm{d}\theta &= \int_0^\pi \frac{(n-1)}{n}\sin^{n-2}(\theta)\,\mathrm{d}\theta - \int_0^\pi \frac{1}{n}\frac{d}{d\theta}\left(\sin^{n-1}(\theta)\cos(\theta)\right)\mathrm{d}\theta \\
&= \frac{(n-1)}{n}\int_0^\pi \sin^{n-2}(\theta)\,\mathrm{d}\theta - \frac{1}{n}\left(\sin^{n-1}(\theta)\cos(\theta)\right)\Big|_0^\pi \\
&= \frac{(n-1)}{n}\int_0^\pi \sin^{n-2}(\theta)\,\mathrm{d}\theta
\end{aligned}$$

E então temos que

$$\begin{aligned}
\int_0^\pi (\sin(\theta))^{2l+1}\,\mathrm{d}\theta &= \frac{2l}{2l+1}\int_0^\pi (\sin(\theta))^{2l-1}\,\mathrm{d}\theta \\
&= \frac{2l}{2l+1}\frac{2l-2}{2l-1}\int_0^\pi (\sin(\theta))^{2l-3}\,\mathrm{d}\theta \\
&= \frac{2l}{2l+1}\frac{2l-2}{2l-1}\frac{2l-4}{2l-3}\int_0^\pi (\sin(\theta))^{2l-5}\,\mathrm{d}\theta
\end{aligned}$$

Iterando várias vezes até chegarmos no $\sin^1(x)$,

$$\int_0^\pi (\sin(\theta))^{2l+1}\,\mathrm{d}\theta = \frac{2l}{2l+1}\frac{2l-2}{2l-1}\frac{2l-4}{2l-3}\cdot\ldots\cdot\frac{2}{3}\int_0^\pi \sin(\theta)\,\mathrm{d}\theta$$

Sabendo que $\int_0^\pi \sin(\theta)\,\mathrm{d}\theta = 2$,

$$\begin{aligned}
\frac{2l}{2l+1}\frac{2l-2}{2l-1}\frac{2l-4}{2l-3}\frac{2l-6}{2l-5}\cdot\ldots\cdot\frac{2}{3} &= \frac{2l}{2l+1}\frac{2(l-1)}{2l-1}\frac{2(l-2)}{2l-3}\frac{2(l-3)}{2l-5}\cdot\ldots\cdot\frac{2}{3} \\
&= \frac{2^l l!}{(2l+1)(2l-1)(2l-3)\cdot\ldots\cdot 3\cdot 1}
\end{aligned}$$

e que

$$(2l+1)(2l-1)(2l-3)\cdot\ldots\cdot 3\cdot 1 = \frac{(2l+1)!}{2l\cdot 2(l-1)\cdot 2(l-2)\cdot\ldots\cdot 2} = \frac{(2l+1)!}{2^l l!}$$

então

$$\int_{-1}^1 (1-x^2)^l\,\mathrm{d}x = \int_0^\pi (\sin(\theta))^{2l+1}\,\mathrm{d}\theta = 2\frac{2^l l!}{\left(\frac{(2l+1)!}{2^l l!}\right)} = \frac{2^{2l+1}(l!)^2}{(2l+1)!}$$

Finalmente, colocando esse resultado na expressão (474), obteremos

$$A_{lm} = \frac{(2l)!}{2^{2l}(l!)^2}\frac{(l+m)!}{(l-m)!}\cdot\frac{2^{2l+1}(l!)^2}{(2l+1)!}$$

$$A_{lm} = \frac{2}{2l+1} \cdot \frac{(l+m)!}{(l-m)!}$$

o que completa a demonstração. ■

I.2 Códigos em Python

I.2.1 Gerador de coeficientes da série de Taylor

```
import numpy as np
import os

dirname = os.path.dirname(__file__)
filename = os.path.join(dirname, 'legendrecoef.txt')

L_MAX = 25            # número de funções de Legendre associados (l = 0 até L_MAX-1)
N_MAX = L_MAX+1       # número de coeficientes da série de taylor
M_MAX = 10           # número de arrays de funções de Legendre (uma para cada m, de m = 0 até m = M_MAX-1)

# a escolha N_MAX = L_MAX + 1  se deve por causa do truncamento que ocorre em l-m. Como m pode ser zero,
# o tamanho máximo do vetor dos coeficientes será ligeramente acima de l por precaução.

salvar = True   # escolha salvar ou não

if(M_MAX > L_MAX):
    ch = input("Aviso: M > L fará algumas funções divergirem em +-1. Manter assim ou limitar M? (y/n)")
    if(ch != "y"):
        M_MAX = L_MAX
    else:
        N_MAX = 200      # alterando número de coeficientes da série de Taylor

A = np.zeros((N_MAX,L_MAX,M_MAX))       # matriz 3D dos coeficientes para cada l e para cada m

def pot(mi,l,parimpar):        # fase de Condon{Shortley + correção
    #fase de Condon{Shortley
    i = 0
    r = 1
    while(i < mi):
        r *= -1
        i += 1

    # correção de sinal
    if(parimpar == 0):
        if((l-mi+2) % 4 == 0):
            r *= -1
    else:
        if((l-mi+1) % 4 ==0):
            r *= -1
```

```
    return r

def calcularcoef():         # usar a relação de recorrência e guardar os resultados no vetor A
    for mi in range(M_MAX):
        print((mi/M_MAX)*100," %")      # progresso
        for li in range(mi,L_MAX,1):    # l >= m
            if((li-mi) % 2 == 0):            # l-m é par
                A[0,li,mi] = pot(mi,li,0)          # a0 = 1 ou -1 (fase de Condon{Shortley)
                A[1,li,mi] = 0          # a1 = 0
                for n in range(0,N_MAX-2,2):        #calcular coeficientes
                    A[n+2,li,mi] = ((n+mi)*(n+mi+1) - li*(li+1))*A[n,li,mi]/((n+2)*(n+1))  # relação de recorrênc
            else:                       # l-m é impar
                A[0,li,mi] = 0          # a0 = 0
                A[1,li,mi] = pot(mi,li,1)          # a1 = 1 ou -1 (fase de Condon{Shortley)
                for n in range(1,N_MAX-2,2):        # calcular coeficientes
                    A[n+2,li,mi] = ((n+mi)*(n+mi+1) - li*(li+1))*A[n,li,mi]/((n+2)*(n+1)) # relação de recorrênci

    if(salvar):
        # salvar os dados em um arquivo de texto
        coef_file = open(filename, "w")
        for row in A:
            np.savetxt(coef_file, row)
        coef_file.close()
    return 1

print("Calculando coeficientes...")
calcularcoef()
print("Pronto!")
```

I.2.2 Gerador das normas das funções associadas

```
import numpy as np
import math
import os

dirname = os.path.dirname(__file__)
filename = os.path.join(dirname, 'legendrenorm.txt')
filenamecoef = os.path.join(dirname, 'legendrecoef.txt')

# nota:
# essas três constantes precisam ser os mesmos do código que gera os coeficientes
# e as que são usadas para fazer gráficos
L_MAX = 25            # número de funções de Legendre associados (l = 0 até L_MAX-1)
N_MAX = L_MAX+1           # número de coeficientes da série de taylor
M_MAX = 10           # número de arrays de funções de Legendre (uma para cada m, de m = 0 até m = M_MAX-1)
```

```
if(M_MAX > L_MAX):
    print("Aviso: M > L terá normas divergindo em +-1. Encerrando algoritmo.")
    exit()

A = np.loadtxt(filenamecoef).reshape(N_MAX, L_MAX,M_MAX)     # ler matriz de coeficientes
C = np.zeros((L_MAX,M_MAX))   # matriz onde ficarão guardadas as normas das funções de Legendre associadas

delta = 10**(-2)            # tamanho da partição da integral (precisão)

def funcao(x,l,m):  # função de legendre associada (série de Taylor vezes o fator (1-x^2)^m/2)
    i = l-m
    soma = 0
    while(i >= 0):
        soma = soma*x + A[i,l,m]
        i = i - 1
    fator = 1-(x*x)
    if(fator*fator >= 0):      # correção de erro de domínio do math.pow
        return math.pow(fator,0.5*m)*soma
    else:
        return 0

def funcao2(x,l,m):      # integrando da norma (não calcular o mesmo valor duas vezes, economizar tempo computacio
    u = funcao(x,l,m)
    return u*u

def gerarnormassimpson():                # calcular as normas usando regra de Simpson
    for mi in range(M_MAX):
        print((mi/M_MAX)*100," %")      # progresso
        for li in range(mi,L_MAX,1):    # l >= m
            integral = 0
            x = -0.99                       # integral de -1 até 1
            while(x <= 0.99):               # regra de Simpson
                integral += (delta/6)*(funcao2(x,li,mi) + 4*funcao2(x+(0.5*delta),li,mi) + funcao2(x+delta,li,mi)
                x += delta
            C[li,mi] = math.sqrt(integral)      # a norma é a raíz quadrada do produto interno da função com ela m
    # salvar os dados em um arquivo de texto
    norm_file = open(filename, "w")
    for row in C:
        np.savetxt(norm_file, row)
    norm_file.close()

print("Gerando normas...")
gerarnormassimpson()
print("Pronto!")
```

I.2.3 Gráficos das funções associadas

```
import numpy as np
import math
import matplotlib.pyplot
import os

# nota:
# essas três constantes precisam ser os mesmos do código que gera os coeficientes
# e as que são usadas para gerar as normas.
L_MAX = 25           # número de funções de Legendre associados (l = 0 até L_MAX-1)
N_MAX = L_MAX+1      # número de coeficientes da série de taylor
M_MAX = 10          # número de arrays de funções de Legendre (uma para cada m, de m = 0 até m = M_MAX-1)

dirname = os.path.dirname(__file__)
filenamecoef = os.path.join(dirname, 'legendrecoef.txt')
filenamenorm = os.path.join(dirname, 'legendrenorm.txt')

A = np.loadtxt(filenamecoef).reshape(N_MAX, L_MAX,M_MAX)     # ler matriz de coeficientes
C = np.loadtxt(filenamenorm).reshape(L_MAX, M_MAX)     # ler matriz das normas

normalizar = True             # normalizar ou não as funções
cossenos = False              # colocar ou não cossenos no argumento das funções
npartes = 3000             # número de cálculos da série de Taylor
m = 1                    # escolha um m
lmax = m + 5             # plotar até qual l
linit = m + 0            # plotar a partir de qual l

# limites do eixo x do gráfico
if(cossenos != 1):
    xi = -1
    xf = 1
else:
    xi = 0
    xf = math.pi

def funcao(x,l,m):  # função de legendre associada (série de Taylor vezes o fator (1-x^2)^m/2)
    i = l-m
    soma = 0
    while(i >= 0):
        soma = soma*x + A[i,l,m]
        i = i - 1
    fator = 1-(x*x)
    if(fator*fator >= 0):      # correção de erro de domínio do math.pow
        return math.pow(fator,0.5*m)*soma
    else:
        return 0
```

```
ex = []
ey = []
ex1 = []
ey1 = []

# plotar eixo x
ex1.append(xi)
ex1.append(xf)
ey1.append(0)
ey1.append(0)

matrizvariavel = np.linspace(xi,xf,npartes)
matrizcalculos = np.zeros((len(matrizvariavel),L_MAX))

if(normalizar == 1):        # normalizar ou não, assim como usar ou não cossenos
    if(cossenos == 0):
        for li in range(linit,lmax,1):
            for i in range(len(matrizvariavel)):
                matrizcalculos[i,li] = funcao(matrizvariavel[i],li,m)/C[li,m]
    else:
        for li in range(linit,lmax,1):
            for i in range(len(matrizvariavel)):
                matrizcalculos[i,li] = funcao(math.cos(matrizvariavel[i]),li,m)/C[li,m]
else:
    if(cossenos == 0):
        for li in range(linit,lmax,1):
            for i in range(len(matrizvariavel)):
                matrizcalculos[i,li] = funcao(matrizvariavel[i],li,m)
    else:
        for li in range(linit,lmax,1):
            for i in range(len(matrizvariavel)):
                matrizcalculos[i,li] = funcao(math.cos(matrizvariavel[i]),li,m)

l = linit
while(l < lmax):        # plotar cada gráfico
    matplotlib.pyplot.plot(matrizvariavel, matrizcalculos[:,l],label='l = ' + str(l))
    l += 1

matplotlib.pyplot.plot(ex1, ey1)
matplotlib.pyplot.title(label='Funções de Legendre associadas com m = ' + str(m))
matplotlib.pyplot.xlabel("x")
matplotlib.pyplot.ylabel("P(x)")
matplotlib.pyplot.legend()
matplotlib.pyplot.show()
```

I.2.4 Gráficos dos harmônicos esféricos

```
from math import *
import matplotlib.pyplot
from numpy import zeros, arctan, arccos
from pylab import imshow,show
import numpy as np
import matplotlib.pyplot as plt
import os

L_MAX = 25           # número de funções de Legendre associados (l = 0 até L_MAX-1)
N_MAX = L_MAX+1      # número de coeficientes da série de taylor
M_MAX = 10          # número de arrays de funções de Legendre (uma para cada m, de m = 0 até m = M_MAX-1)

dirname = os.path.dirname(__file__)
filenamecoef = os.path.join(dirname, 'legendrecoef.txt')
filenamenorm = os.path.join(dirname, 'legendrenorm.txt')

A = np.loadtxt(filenamecoef).reshape(N_MAX, L_MAX,M_MAX)     # ler matriz de coeficientes
C = np.loadtxt(filenamenorm).reshape(L_MAX, M_MAX)     # ler matriz das normas das funções de Legendre

npartes = 150          # número de linhas e colunas, ou seja, haverão npartes ao quadrado de pontos
m = 1                  # escolha um m
l = 1                  # escolha um l

if(m > l):
    print("Erro: M > L")
    exit()

xi = 0      # começar a calcular as funções associadas de Legendre normalizadas com cossenos
xf = pi
delta = (xf - xi)/npartes    # distância entre variáveis para computar a função
normalizacao = 1/sqrt(2*pi)     # constante para normalizar a parte azimutal

def legendre(x,l,m):  # função de legendre associada (série de Taylor vezes o fator (1-x^2)^m/2)
    i = l-m
    soma = 0
    while(i >= 0):
        soma = soma*x + A[i,l,m]
        i = i - 1
    fator = 1-(x*x)
    if(fator*fator >= 0):      # correção de erro de domínio do math.pow
        return math.pow(fator,0.5*m)*soma
    else:
        return 0

n = npartes  # quantidade de pontos por linha/coluna
```

```
thetas = np.linspace(0.01,pi,npartes)        # valores de theta, 0 a pi
phis = np.linspace(0,2*pi,npartes)    # valores de phi, 0 a 2*pi
harmonicos = np.zeros((len(thetas),len(phis)))  # valores de Ylm

for theta in range(len(thetas)):    # calcular o Ylm em todos os pontos da esfera
    for phi in range(len(phis)):
        harmonicos[theta,phi] = normalizacao*(cos(m*phis[phi]))*legendre(cos(thetas[theta]),l,m)/C[l,m]

matrizvelha = zeros([n,n], dtype = float)

# limites do plot do gráfico (e do domínio a ser usado)
# eixo x: phi   /   eixo y: theta
left = 0
right = 2*pi
bot = 0
top = pi

def harmonic(x,y):   # converter x e y do gráfico nos parâmetros de theta e phi
    i = int(np.floor(npartes*(y/pi)-1))
    j = int(np.floor(npartes*(x/(2*pi))-1))
    return harmonicos[i,j]

# função de duas variáveis
def funcao(x,y):
    return harmonic(x,y)

# função de duas variáveis em coordenadas polares (não usado)
def pfuncao(r,theta):
    return theta

# calcular theta (se for preciso)
# obs: theta das coordenadas polares do gráfico
def artheta2(x,y,r):
    if(r != 0):
        if(y >= 0):
            return arccos(x/r)
        if(x < 0):
            return pi + arctan(y/x)
        if(x == 0):
            return 3*pi*0.5
        return pi + pi + arctan(y/x)
    return 0

i=0
j=0
while(i < n):
    x = (right-left)*i/(n-1) + left                 # calcular o 'x' em termos do índice i
```

```
    while(j < n):
        y = (bot-top)*j/(n-1) + top                 # calcular o 'y' em termos do índice j
        r = sqrt(x*x+y*y)                           # calcular o r em coordenadas polares
        theta = artheta2(x,y,r)                        # calcular o theta em coordenadas polares
        matrizvelha[j,i] = funcao(x,y)              # função de duas variáveis em coordenadas cartesianas
        #matrizvelha[j,i] = pfuncao(r,theta)        # função de duas variáveis em coordenadas polares
        # obs: as coordenadas polares são em relação ao plano do gráfico, e não às grandezas dos parâmetros usados
        # Por isso estamos usando coordenadas cartesianas.
        j += 1
    i += 1
    j = 0

matplotlib.pyplot.xlabel("Phi")
matplotlib.pyplot.ylabel("Theta")
matplotlib.pyplot.title(label='Harmônicos esféricos com m = ' + str(m) + ' e l = ' + str(l))
imshow(matrizvelha, cmap='jet', extent=[left, right, bot, top])  # imprimir gráfico
plt.ylim(pi, 0)   # inverter eixo y, assim a parte de cima é theta = 0

eixocor = plt.axes([0.10, 0.89, 0.8, 0.06])             # barra de cores
plt.colorbar(cax=eixocor,orientation = 'horizontal')
show()
```

J Polinômios associados de Laguerre

J.1 Códigos em Python

J.1.1 Gerador de coeficientes

```
import numpy as np
import math
import os

POL_MAX = 25            # número de polinômios de Laguerre associados
N_MAX = POL_MAX+1       # número de coeficientes da série de taylor
L_MAX = 2               # relacionado ao índice superior do polinômio de Laguerre associado (para o átomo de hidro

dirname = os.path.dirname(__file__)
filenamecoef = os.path.join(dirname, 'lagcoef.txt')

A = np.zeros((N_MAX,POL_MAX,L_MAX+1))       # matriz 3D dos coeficientes para cada polinômio e para cada l em 2l+1

def calcularcoef():         # usar a relação de recorrência e guardar os resultados no vetor A
    for npol in range(POL_MAX):
        for li in range(L_MAX+1):
            A[0,npol,li] = a0(npol,li)
            for n in range(N_MAX-1):
```

```
                A[n+1,npol,li] = 2*(n-npol)*A[n,npol,li]/((n+1)*(n + 2*li + 2))     # relação de recorrência do á
                #print("pol = ", npol, " / l = ", li, " / coef ",n+1," = ", A[n+1,npol,li])
                if (A[n+1,npol,li] == 0):
                    break

    # salvar os dados em um arquivo de texto
    coef_file = open(filenamecoef, "w")
    for row in A:
        np.savetxt(coef_file, row)
    coef_file.close()
    return 1

def a0(npol,li):        # calcular constante inicial
    prod = 1
    for k in range(npol):
        prod *= li + k + 1
    return prod/math.factorial(npol)

def funcao(x,nl,li):              # Polinômio do átomo de hidrogênio (Laguerre modificado)
    soma = 0
    for n in range(N_MAX-1):
        soma += A[n,nl,li]*math.pow(x,n)
    return soma


print("Calculando coeficientes...")
calcularcoef()
print("Pronto!")
```

J.1.2 Gerador das normas

```
import numpy as np
import math
import os

POL_MAX = 25            # número de polinômios de Laguerre associados
N_MAX = POL_MAX+1       # número de coeficientes da série de taylor
L_MAX = 2               # relacionado ao índice superior do polinômio de Laguerre associado (para o átomo de hidro

dirname = os.path.dirname(__file__)
filenamecoef = os.path.join(dirname, 'lagcoef.txt')
filenamenorm = os.path.join(dirname, 'lagnorm.txt')

A = np.loadtxt(filenamecoef).reshape(N_MAX,POL_MAX,L_MAX+1)    # ler matriz dos coeficientes
C = np.zeros((POL_MAX,L_MAX+1))     # matriz das normas para cada l
```

```
delta = 10**(-2)        # precisão da integral
xf = 30                 # número muito alto, integral infinita

def funcao(x,nl,li):  # Polinômio do átomo de hidrogênio (Laguerre modificado)
    i = N_MAX - 1
    soma = A[i,nl,li]*x + A[i-1,nl,li]
    i = i-2
    while(i >= 0):
        soma = soma*x + A[i,nl,li]
        i = i-1
    return soma

def funcao2(x,npol,li):             # o integrando do cálculo da norma
    u = math.pow(x,li)*math.exp(-x)*funcao(x,npol,li)
    return x*u*u

def gerarnormas():              # cálcular a intergal das normas usando regra de Simpson
    for nl in range(POL_MAX):
        for li in range(L_MAX+1):
            integral = 0
            x = 0
            while(xf >= x):
                integral += (delta/6)*(funcao2(x,nl,li) + 4*funcao2(x+(0.5*delta),nl,li) + funcao2(x+delta,nl,li)
                x += delta
            C[nl,li] = math.sqrt(integral)
        print(100*(nl/POL_MAX),"%")
    # salvar dados em um txt
    norm_file = open(filenamenorm, "w")
    for row in C:
        np.savetxt(norm_file, row)
    norm_file.close()

print("Gerando normas...")
gerarnormas()
print("Pronto!")
```

J.1.3 Expansão em autofunções de Sturm-Liouville

```
import numpy as np
import math
import matplotlib.pyplot
import os

POL_MAX = 25            # número de polinômios de Laguerre associados
N_MAX = POL_MAX+1       # número de coeficientes da série de taylor
L_MAX = 2               # relacionado ao índice superior do polinômio de Laguerre associado (para o átomo de hidro
```

```python
dirname = os.path.dirname(__file__)
filenamecoef = os.path.join(dirname, 'lagcoef.txt')
filenamenorm = os.path.join(dirname, 'lagnorm.txt')

A = np.loadtxt(filenamecoef).reshape(N_MAX,POL_MAX,L_MAX+1) # ler matriz dos coeficientes
C = np.loadtxt(filenamenorm).reshape(POL_MAX,L_MAX+1)   # ler matriz das normas

D = np.zeros(POL_MAX)     # vetor dos coeficientes da expansão em autofunções
l = 2                    # escolha um l
delta = 10**(-2)         # precisão da integral
xf = 60                  # limite do gráfico no eixo x

ncoeficientes = 25     # número de coeficientes da expansão em autofunções

if(ncoeficientes > D.size):
    raise Exception("Não há polinômios suficientes!")
if(l > L_MAX):
    raise Exception("modos 'l' insuficientes")

def funcao(x,nl,li):  # Polinômio do átomo de hidrogênio (Laguerre modificado)
    i = N_MAX - 1
    soma = A[i,nl,li]*x + A[i-1,nl,li]
    i = i-2
    while(i >= 0):
        soma = soma*x + A[i,nl,li]
        i = i-1
    return soma

def g(x):                          # qualquer função quadrado-integrável em [0, +infinito)
    return math.sin(x)

def funcintegra(x,nl,li):          # o integrando do cálculo das coordenadas
    return g(x)*math.pow(x,li+1)*math.exp(-x)*funcao(x,nl,li)/C[nl,li]

def calcularcoord():               # calcular as coordenadas da expansão
    for npol in range(POL_MAX):
        integral = 0
        x = 0
        while(xf >= x):
            integral += (delta/6)*(funcintegra(x,npol,l) + 4*funcintegra(x+(0.5*delta),npol,l) + funcintegra(x+de
            x += delta
        D[npol] = integral


ex = []
ey = []
ex1 = []
```

```
ey1 = []
ey2 = []

x = 0               # inicio
npartes = 1000      # número de partições
delta2 = (xf-x)/npartes     # tamanho das partições

# imprimir o eixo x
ex1.append(0)
ex1.append(xf)
ey1.append(0)
ey1.append(0)

def expansao(x):       # calcular a expansão em autofunções
    soma = 0
    for npol in range(ncoeficientes):
        soma += D[npol]*math.pow(x,l)*math.exp(-x)*funcao(x,npol,l)/C[npol,l]
    return soma

print("Calculando coordenadas...")
calcularcoord()
print("Calculando os valores da expansão...")

while(x <= xf):     # calculando a expansão para cada 'x'
    ex.append(x)
    ey.append(expansao(x))
    ey2.append(g(x))
    x += delta2
print("Pronto!")

matplotlib.pyplot.plot(ex, ey, label='exp') # label='l = 0'
matplotlib.pyplot.plot(ex1, ey1)
matplotlib.pyplot.plot(ex, ey2, label='g(x)')
matplotlib.pyplot.title(label='Expansão em autofunções do átomo de hidrogênio com l = ' + str(l))
matplotlib.pyplot.xlabel("x")
matplotlib.pyplot.ylabel("g(x)")
matplotlib.pyplot.legend()
matplotlib.pyplot.show()
```

K Identidade de Abel

Teorema K.1 (Identidade de Abel)**.** *Dada a equação diferencial*

$$f''(x) + p(x)f'(x) + q(x)f(x) = 0$$

com $p(x)$ e $q(x)$ funções contínuas em um intervalo I. Então o wronskiano das duas soluções linearmente independentes é dado pela expressão

$$W(x) = C \cdot \exp\left(-\int_{x_0}^{x} p(x')\,\mathrm{d}x'\right) \quad \forall x \in I$$

Demonstração. Sejam duas funções $f_1(x)$ e $f_2(x)$ que satisfazem

$$f_1''(x) + p(x)f_1'(x) + q(x)f_1(x) = 0 \tag{475}$$

$$f_2''(x) + p(x)f_2'(x) + q(x)f_2(x) = 0 \tag{476}$$

Multiplicando (475) por $f_2(x)$, a equação (476) por $f_1(x)$ e subtraindo uma equação da outra, o termo proporcional a $q(x)$ desaparece e ficamos com

$$[f_2''(x)f_1(x) - f_2(x)f_1''(x)] + p(x) \cdot [f_2'(x)f_1(x) - f_1'(x)f_2(x)] = 0 \tag{477}$$

Note que o segundo termo entre parênteses é o determinante da matriz wronskiana

$$W(x) = \det \begin{bmatrix} f_1(x) & f_2(x) \\ f_1'(x) & f_2'(x) \end{bmatrix}$$

enquanto que o primeiro termo entre parênteses é o determinante da matriz

$$\begin{bmatrix} f_1(x) & f_2(x) \\ f_1''(x) & f_2''(x) \end{bmatrix} \tag{478}$$

A matriz (478) é obtida derivando a última linha da matriz wronskiana e portanto, de acordo com o teorema (1.41), é a derivada da matriz wronskiana. Logo o primeiro termo entre parênteses é $W'(x)$.

Então a equação (477) fica

$$W'(x) + p(x)W(x) = 0$$

e então

$$\frac{W'(x)}{W(x)} = -p(x)$$

O primeiro membro é a derivada de $\ln(W(x)) + D$, com D constante e portanto

$$\frac{d}{dx}\left(\ln(W(x)) + D\right) = -p(x)$$

Integrando ambos os membros e usando o teorema fundamental do cálculo, obtemos

$$\ln(W(x)) + D = -\int_{x_0}^{x} p(x')\,\mathrm{d}x'$$

e então, $\forall x \in I$ temos

$$W(x) = C \cdot \exp\left(-\int_{x_0}^{x} p(x')\,\mathrm{d}x'\right)$$

onde $C = e^{-D}$.

■

Referências

[1] T. M. APOSTOL, *Cálculo: Cálculo com funções de uma variável, com introdução à álgebra linear*, vol. 1. EDITORIAL REVERTÉ, S. A, 2 ed., 1988.

[2] D. G. FIGUEIREDO, *Análise de Fourier e Equações Diferenciais Parciais*. Projeto Euclides, IMPA, 3 ed., 1977.

[3] N. A. LEMOS, *Mecânica Analítica*. Livraria da Física, 2 ed., 2004.

[4] D. J. GRIFFITHS, *Mecânica Quântica*. PEARSON, 2 ed., 2011.

[5] T. M. APOSTOL, *Cálculo: funções de várias variáveis, álgebra linear e aplicações às equações diferenciais e às probabilidades*, vol. 2. EDITORIAL REVERTÉ, S. A, 2 ed., 1996.

[6] L. I. SCHIFF, *Quantum Mechanics*. McGRAW-HILL, 1 ed., 1949.

[7] E. L. LIMA, *Curso de Análise*, vol. 1. Instituto de Matemática Pura e Aplicada, Projeto Euclides, 7 ed., 1976.

[8] M. N. OLIVEIRA, "O teorema de arzelá-ascoli e aplicações," trabalho de conclusão de curso, Universidade Estadual da Paraíba, Março 2014.

[9] M. A. AL-GWAIZ, *Sturm-Liouville Theory and its Applications*. Springer, 2008.

[10] R. B. Guenther and J. W. Lee, *Sturm-Liouville Problems: Theory and Numerical Implementation*. CRC Press, 2019.